

**(12) STANDARD PATENT**  
**(19) AUSTRALIAN PATENT OFFICE**

(11) Application No. **AU 2017245123 B2**

(54) Title  
**Carotenoid production method**

(51) International Patent Classification(s)  
**C12N 15/09** (2006.01) **C12N 5/10** (2006.01)  
**C12N 1/15** (2006.01) **C12P 23/00** (2006.01)  
**C12N 1/19** (2006.01) **C12Q 1/02** (2006.01)  
**C12N 1/21** (2006.01)

(21) Application No: **2017245123** (22) Date of Filing: **2017.03.29**

(87) WIPO No: **WO17/171098**

(30) Priority Data

(31)	Number	(32)	Date	(33)	Country
	<b>2016-071303</b>		<b>2016.03.31</b>		<b>JP</b>

(43) Publication Date: **2017.10.05**

(44) Accepted Journal Date: **2023.05.25**

(71) Applicant(s)  
**JXTG Nippon Oil & Energy Corporation**

(72) Inventor(s)  
**Nagai, Hidetada;Sato, Wataru;Takahashi, Toshiyuki;Sato, Harumi**

(74) Agent / Attorney  
**FB Rice Pty Ltd, L 33 477 Collins Street, Melbourne, VIC, 3000, AU**

(56) Related Art  
**US 2014/0148622 A1**  
**WO 2015/189428 A1**  
**WO 2006/018211 A1**

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関  
国際事務局

(43) 国際公開日  
2017 年 10 月 5 日(05.10.2017)



(10) 国際公開番号

WO 2017/171098 A1

- (51) 国際特許分類:  
*C12N 15/09* (2006.01) *C12N 5/10* (2006.01)  
*C12N 1/15* (2006.01) *C12P 23/00* (2006.01)  
*C12N 1/19* (2006.01) *C12Q 1/02* (2006.01)  
*C12N 1/21* (2006.01)
- (21) 国際出願番号: PCT/JP2017/014162
- (22) 国際出願日: 2017 年 3 月 29 日(29.03.2017)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:  
特願 2016-071303 2016 年 3 月 31 日(31.03.2016) JP
- (71) 出願人: J X エネルギー株式会社 (JX NIPPON OIL & ENERGY CORPORATION) [JP/JP]; 〒1008162 東京都千代田区大手町一丁目 1 番 2 号 Tokyo (JP).
- (72) 発明者: 永井 秀忠(NAGAI, Hidetada); 〒1008162 東京都千代田区大手町一丁目 1 番 2 号 J X エネルギー株式会社内 Tokyo (JP). 佐藤 渉(SATO, Wataru); 〒1008162 東京都千代田区大手町一丁目 1 番 2 号 J X エネルギー株式会社内 Tokyo (JP). 高橋 季之(TAKAHASHI, Toshiyuki); 〒1008162 東京都千代田区大手町一丁目 1 番 2 号 J X エネルギー株式会社内 Tokyo (JP). 佐藤 温美(SATO, Harumi); 〒1008162 東京都千代田区大手町一丁目 1 番 2 号 J X エネルギー株式会社内 Tokyo (JP).
- (74) 代理人: 小林 浩, 外(KOBAYASHI Hiroshi et al.); 〒1040028 東京都中央区八重洲二丁目 8 番 7 号

福岡ビル 9 階 阿部・井窪・片山法律事務所  
Tokyo (JP).

- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

- 国際調査報告 (条約第 21 条(3))
- 規則 13 の 2 に基づいて明細書とは別に提出された、寄託された生物材料に関する表示 (規則 13 の 2.4(d)(i) 及び 48.2(a)(viii))
- 明細書の別個の部分として表した配列リスト (規則 5.2(a))

(54) Title: CAROTENOID PRODUCTION METHOD

(54) 発明の名称: カロテノイドの製造方法

(57) Abstract: A mutant carotenoid-producing bacterium that includes a gene set forth in any of (a) to (c). (a) A gene of a carotenoid-producing bacterium, which codes for a protein that includes a mutant amino acid sequence in which at least the 225th amino acid residue of an amino acid sequence of 1-deoxy-D-xylulose 5-phosphate synthase is substituted with another amino acid residue; (b) a gene of a carotenoid-producing bacterium, which codes for a protein that includes a mutant amino acid sequence in which at least the 305th amino acid residue of an amino acid sequence of decaprenyl diphosphate synthase is substituted with another amino acid residue; and (c) both of the genes in (a) and (b).

(57) 要約: 以下の (a) ~ (c) のいずれかの遺伝子を含む、変異型カロテノイド産生細菌。 (a) カロテノイド産生細菌における 1-デオキシ-D-キシロース 5リン酸合成酵素のアミノ酸配列において、少なくとも第 225 番目のアミノ酸残基が他のアミノ酸残基に置換された変異型アミノ酸配列を含むタンパク質をコードする遺伝子 (b) カロテノイド産生細菌におけるデカプレニルニリン酸合成酵素のアミノ酸配列において、少なくとも第 305 番目のアミノ酸残基が他のアミノ酸残基に置換された変異型アミノ酸配列を含むタンパク質をコードする遺伝子 (c) 上記 (a) 及び (b) の両方の遺伝子



WO 2017/171098 A1

## DESCRIPTION

### CAROTENOID PRODUCTION METHOD

#### FIELD OF THE INVENTION

[0001] The present invention relates to a method for producing a carotenoid using a mutant strain of a carotenoidogenic bacterium.

#### BACKGROUND ART

[0002] Carotenoids are useful natural pigments that can be used as feed additives, food additives, pharmaceutical products and the like. Examples of carotenoids include astaxanthin, canthaxanthin, zeaxanthin,  $\beta$ -cryptoxanthin, lycopene,  $\beta$ -carotene, adonirubin, adonixanthin, echinenone, asteroidenone and 3-hydroxyechinenone. Among them, astaxanthin is useful as a feed additive such as a body color improving agent for farmed fish such as salmon, trout and red sea bream, and an egg-yolk color improving agent for poultry and the like. Natural astaxanthins are also highly valuable in industries as safe food additives and health food ingredients. Similar to astaxanthin, adonixanthin and adonirubin are expected of their use as feed additives, food additives, pharmaceutical products and the like.

[0003] In addition,  $\beta$ -carotene is used as a feed additive, a food additive, a pharmaceutical product and a like, canthaxanthin is used as a feed additive, a food additive, a cosmetic product and the like, and zeaxanthin is used as a food additive, a feed additive and the like. Furthermore, lycopene, echinenone,  $\beta$ -cryptoxanthin, 3-hydroxyechinenone and asteroidenone are also expected of their use as feed additives, food ingredients and the like. As a method for producing such carotenoids, a chemical synthesis method, an extraction method from a natural product, a production method by culturing microorganisms, and the like are known.

[0004] As the chemical synthesis method for astaxanthin, a method by  $\beta$ -carotene conversion (Pure Appl. Chem., 57, 741, 1985 (Non-patent document 1)) and a method by synthesizing from a C15 phosphonium salt (Helv. Chim. Acta, 64, 2436, 1981 (Non-patent

document 2)) are known. As the extraction method from a natural product, astaxanthin may be extracted and collected from fish like salmon or red sea bream, or a crustacean like shrimp, crab or krill in which astaxanthin is present.

[0005] Examples of the method for producing a carotenoid using microorganisms include a culture method using a green alga *Haematococcus pluvialis* (Japanese Patent Application Publication No. 2007-97584 (Patent document 1)), a fermentation method using red yeast *Phaffia rhodozyma* (Japanese Patent Application Publication No. H11(1999)-69969 (Patent document 2)), a fermentation method using a bacterium belonging to the genus *Paracoccus* (hereinafter, also referred to as a “bacterium of the genus *Paracoccus*”), a fermentation method using a bacterium belonging to the genus *Brevundimonas* (Japanese Patent Application Publication No. 2006-340676 (Patent document 3)), and a fermentation method using a bacterium belonging to the genus *Erythrobacter* (Japanese Patent Application Publication No. 2008-259449 (Patent document 4)). Examples of the bacteria of the genus *Paracoccus* that produce carotenoids include strains E-396 and A-581-1 (Japanese Patent Application Publication No. H7(1995)-79796 (Patent document 5) and International Journal of Systematic Bacteriology (1999), 49, 277-282 (Non-patent document 3)). Examples of other carotenoid-producing bacteria belonging to the genus *Paracoccus* include *Paracoccus marcusii* strain MH1 (Japanese Patent Application Publication No. 2001-512030 (Patent document 6)), *Paracoccus haeundaensis* strain BC74171 (International Journal of Systematic and Evolutionary Microbiology (2004), 54, 1699-1702 (Non-patent document 4)), bacterium *Paracoccus* strain sp. N-81106 (Japanese Patent Application Publication No. 2007-244205 (Patent document 7)), *Paracoccus zeaxanthinifaciens* (International Journal of Systematic and Evolutionary Microbiology (2003), 53, 231-238 (Non-patent document 5)) and *Paracoccus* sp. strain PC-1 (pamphlet of WO2005/118812 (Patent document 8)).

[0006] The above-mentioned methods for producing carotenoids, however, have several problems. For example, a carotenoid produced by a chemical synthesis method may be safe but gives unfavorable impression to the consumers. The production cost of carotenoids extracted from natural products are much more expensive than the chemical synthesis method. Among the productions using microorganisms, production by culturing a gree alga or a yeast is low in productivity and difficult in extracting carotenoids from the cultured products since these microorganisms have strong cell walls.



[0007] Meanwhile, several culture methods and production methods have been reported for carotenoid productions using bacteria belonging to the genus *Paracoccus* because they are advantageous in terms of fast bacterial growth rate, high carotenoid productivity, and easy carotenoid extraction from the cultured product.

[0008] For example, Japanese Patent Application Publication No. 2007-143492 (Patent document 9) discloses a method of adding an iron salt during culturing, pamphlet of WO2010/044469 (Patent document 10) discloses a method of adding amino acids to the medium, Japanese Patent Application Publication No. 2011-188795 (Patent document 11) discloses a method of adding biotin to the medium, and Japanese Patent Application Publication No. 2012-139164 (Patent document 12) discloses a method of adding a calcium compound to the medium to 3.6 mM or more.

However, details as to which gene of the carotenoid-producing bacteria contributes to the increase in the production efficiency have been unclear.

#### PRIOR ART DOCUMENTS

[0009] Patent documents

Patent document 1: Japanese Patent Application Publication No. 2007-97584

Patent document 2: Japanese Patent Application Publication No. H11(1999)  
-69969

Patent document 3: Japanese Patent Application Publication No. 2006-340676

Patent document 4: Japanese Patent Application Publication No. 2008-259449

Patent document 5: Japanese Patent Application Publication No. H7(1995)-79796

Patent document 6: Japanese Patent Application Publication No. 2001-512030

Patent document 7: Japanese Patent Application Publication No. 2007-244205

Patent document 8: Pamphlet of WO2005/118812

Patent document 9: Japanese Patent Application Publication No. 2007-143492

Patent document 10: Pamphlet of WO2010/044469

Patent document 11: Japanese Patent Application Publication No. 2011-188795

Patent document 12: Japanese Patent Application Publication No. 2012-139164

[0010] Non-patent documents

Non-patent document 1: Pure Appl. Chem., 57, 741, 1985

Non-patent document 2: Helv. Chim. Acta, 64, 2436, 1981

Non-patent document 3: International Journal of Systematic Bacteriology (1999), 49, 277-282

Non-patent document 4: International Journal of Systematic and Evolutionary Microbiology (2004), 54, 1699-1702

Non-patent document 5: International Journal of Systematic and Evolutionary Microbiology (2003), 53, 231-238

## SUMMARY OF INVENTION

### Problem to be Solved by Invention

[0011] The present invention has an objective of providing a mutant carotenoidogenic bacterium, and a method for producing a carotenoid using said bacterium.

[0012] The present inventors have gone through extensive investigation to solve the above-described problems, and as a result of which succeeded in acquiring a bacterium having high astaxanthin productivity among the bacteria subjected to a mutation treatment, thereby accomplishing the present invention.

### Means for Solving Problem

[0013] (1) A mutant carotenoidogenic bacterium, comprising any of genes (a)-(c) below:

(a) a gene encoding a protein comprising a mutant amino acid sequence in which at least the 225th amino acid residue in the amino acid sequence of 1-deoxy-D-xylulose 5-phosphate synthase of a carotenoidogenic bacterium has been substituted with other amino acid residue;

(b) a gene encoding a protein comprising a mutant amino acid sequence in which at least the 305th amino acid residue in the amino acid sequence of decaprenyl diphosphate synthase of a carotenoidogenic bacterium has been substituted with other amino acid residue; and

(c) both of the genes (a) and (b) above.

(2) The bacterium according to (1), wherein the amino acid sequence of 1-deoxy-D-xylulose 5-phosphate synthase is the sequence represented by SEQ ID NO:2.

(3) The bacterium according to either one of (1) and (2), wherein the 225th amino acid residue has been substituted from glycine to aspartic acid.

- (4) The bacterium according to any one of (1)-(3), wherein the amino acid sequence of decaprenyl diphosphate synthase is the sequence represented by SEQ ID NO:4.
- (5) The bacterium according to any one of (1)-(4), wherein the 305th amino acid residue has been substituted from alanine to valine.
- (6) The bacterium according to any one of (1)-(5), which has acquired carotenogenic capacity that is higher than the carotenogenic capacity of a carotenoidogenic bacterium without the gene encoding the protein comprising the mutant amino acid sequence.
- (7) The bacterium according to (6), which has acquired carotenogenic capacity that is at least 5 times or more the carotenoid production amount of a carotenoidogenic bacterium without the gene encoding the protein comprising the mutant amino acid sequence.
- (8) The bacterium according to any one of (1)-(7), wherein the carotenoidogenic bacterium belongs to the genus *Paracoccus*.
- (9) The bacterium according to (8), wherein the bacterium belonging to the genus *Paracoccus* is strain E-396.
- (10) The bacterium according to any one of (1)-(9), wherein the carotenoid is astaxanthin.
- (11) A method for producing a carotenoid, comprising culturing the bacterium according to any one of (1)-(10), and collecting the carotenoid from the resulting cultured product.
- (12) The method according to (11), wherein the carotenoid production amount is at least 5 times or more the carotenoid production amount of a carotenoidogenic bacterium without the gene encoding the protein comprising the mutant amino acid sequence.
- (13) The method according to either one of (11) and (12), wherein the carotenoid is astaxanthin.
- (14) A method for screening for a carotenoidogenic bacterium, comprising subjecting a carotenoidogenic bacterium to a mutation treatment, and selecting a bacterium having any of characteristics (a)-(c) below from the bacteria subjected to the mutation treatment:
- (a) a characteristic where the activity of 1-deoxy-D-xylulose 5-phosphate synthase is increased compared to said activity in the bacterium before the mutation treatment;
  - (b) a characteristic where the activity of decaprenyl diphosphate synthase is decreased compared to said activity in the bacterium before the mutation treatment; and

- (c) both of the characteristics (a) and (b) above.
- (15) A method for producing a carotenoid, comprising culturing the bacterium selected by the method according to (14), and collecting a carotenoid from the resulting cultured product.
- (16) A gene encoding a protein comprising a mutant amino acid sequence in which at least the 225th amino acid residue in the amino acid sequence of 1-deoxy-D-xylulose 5-phosphate synthase has been substituted with other amino acid residue.
- (17) A gene comprising either of DNA (a) or (b) below:
- (a) DNA comprising the nucleotide sequence represented by SEQ ID NO:5; or
  - (b) DNA that hybridizes with DNA having a nucleotide sequence complementary to said DNA (a) under stringent conditions, and that encodes a protein having 1-deoxy-D-xylulose 5-phosphate synthase activity.
- (18) A gene encoding a protein comprising a mutant amino acid sequence in which at least the 305th amino acid residue in the amino acid sequence of decaprenyl diphosphate synthase has been substituted with other amino acid residue.
- (19) A gene comprising either of DNA (a) or (b) below:
- (a) DNA comprising the nucleotide sequence represented by SEQ ID NO:7; or
  - (b) DNA that hybridizes with DNA having a nucleotide sequence complementary to said DNA (a) under stringent conditions, and that encodes a protein with decreased decaprenyl diphosphate synthase activity.
- (20) A recombinant vector comprising any of genes (a)-(c) below:
- (a) the gene according to (16) or (17);
  - (b) the gene according to (18) or (19); and
  - (c) the genes (a) and (b) above.
- (21) A transformant comprising the recombinant vector according to (20).
- (22) A method for producing a carotenoid, comprising culturing the transformant according to (21), and collecting a carotenoid from the resulting cultured product.

#### EFFECT OF THE INVENTION

[0014] The present invention provides a high carotenoid producing bacterium. Carotenoids can efficiently be produced by using the bacterium of the present invention.

## BRIEF DESCRIPTION OF THE DRAWINGS

[0015] [Figure 1] A diagram showing specific productivities of total carotenoid and astaxanthin in strain E-396 and strain ASB-57.

[Figure 2] An image showing a template structure of Enzyme A.

[Figure 3A] A depiction showing the alignment between Enzyme A and the template structure (2O1X). The deduced active sites are depicted in green. The disordered region is represented by a blue bar.

[Figure 3B] A depiction showing the alignment between Enzyme A and the template structure (2O1X). The deduced active sites are depicted in green. The disordered region is represented by a blue bar.

[Figure 4] Images showing the template structure and a model structure of a complex of Enzyme A with TPP coenzyme. The structure of the complex of *D. radiodurans*-derived 1-deoxy-D-xylulose 5-phosphate synthase (DXS) with TPP (2O1X) as the template (left) and the model structure of constructed Enzyme A (right). In the formed homodimer, each subunit binds TPP and Mg. TPP and Mg are shown in space-filling representation.

[Figure 5] Images showing the model structure of the complex of Enzyme A with a TPP intermediate and a substrate. A complex model in which the constructed Enzyme A is bonded with a hydroxyethyl-TPP intermediate having pyruvic acid attached to TPP coenzyme and glyceraldehyde 3-phosphate (GAP) is shown.

[Figure 6] A depiction showing the interaction between Enzyme A and TPP coenzyme.

[Figure 7] Images showing the amino acid residues of Enzyme A that were predicted responsible for the interaction with TPP. The residues predicted to be responsible for the interaction with TPP are represented by sticks.

[Figure 8] Images showing the amino acid residues that were predicted responsible for the interaction between Enzyme A and the hydroxyethyl group of the hydroxyethyl-TPP intermediate. The residues predicted to be responsible for the interaction with the hydroxyethyl group are represented by sticks.

[Figure 9] Images showing the amino acid residues that were predicted

responsible for the interaction between Enzyme A and glyceraldehyde 3-phosphate (GAP).

[Figure 10] A depiction showing the interaction between Enzyme A and the substrate pyruvic acid.

[Figure 11] A depiction showing the interaction between Enzyme A and the substrate glyceraldehyde 3-phosphate.

[Figure 12] A depiction showing the alignment between Enzyme A and other types of DXS. EnzymeA: Enzyme A, DXS\_ECOLI: DXS (*E.coli*), DXS\_VITVI: DXS (*Vitis vinifera*), DXS\_DEIRA: (*Deinococcus radiodurans*). The disordered region in the template structure (DXS\_DEIRA) is represented by a blue bar. The active sites are shown in green boxes while mutations that showed activity enhancement in DXS\_ECOLI and DXS\_VITVI are marked with red triangles.

[Figure 13] Images showing the location of the disordered region in Enzyme A. Asn180 and Met182 in the loop on the N-terminal side of the disordered region (dotted blue line) bind Mg while Ile184 binds TPP.

[Figure 14] Images showing a model structure of the disordered region of Enzyme A. The model structure of the disordered region (light blue) in Enzyme A (left) and the electrostatic potential maps of the model structures of the disordered regions (right).

[Figure 15] A diagram showing the deduced effects of mutant enzyme A in the astaxanthin synthesis pathway.

[Figure 16] An image showing a template structure of Enzyme C. *Rhodobacter capsulatus*-derived decaprenyl diphosphate synthase (PDB ID: 3MZV) is shown.

[Figure 17] A depiction showing the alignment between Enzyme C and the template structure (3MZV). The deduced active sites are depicted in green.

[Figure 18] Images showing the template structure and a model structure of a complex of Enzyme C with IPP and FPP. *R. capsulatus*-derived decaprenyl diphosphate synthase (3MZV) as the template (left) and the model structure of constructed Enzyme C (right). FPP and IPP are shown in space-filling representation.

[Figure 19] Images showing the template structure, and the model structure of the complex of Enzyme C with IPP and FPP. FPP and IPP bind each other in the head-to-tail direction while the condensation reaction occurs between an isopentenyl group of IPP and a phosphate group of FPP (left image, arrow). In the long-chain prenyl diphosphate synthase,

the reaction product further binds IPP and extends deep into the substrate binding site (right image, arrow).

[Figure 20] Images showing comparison between the template structure and the model structure of Enzyme C. The complex models of template *R. capsulatus*-derived decaprenyl diphosphate synthase (left) and Enzyme C (right) with the substrate. The structures with matching amino acid residues are depicted in green. The substrate binding region and its surrounding structure match entirely.

[Figure 21] A depiction showing the alignment between Enzyme C and decaprenyl diphosphate synthase (*Paracoccus zeaxanthinifaciens*).

[Figure 22] Images showing a complex model of Enzyme C with FPP and IPP. The ribbon representation (left) and the surface profile (right). Chain A (light red) and chain B (light blue) are shown. FPP and IPP are shown in space-filling representation.

[Figure 23] Images showing the amino acid residues of Enzyme C predicted to be responsible for the interactions with FPP (top), IPP (center) and Mg (bottom).

[Figure 24] A depiction showing the interaction of Enzyme A with substrates FPP and IPP.

[Figure 25] A depiction showing the interaction of Enzyme A with Mg.

[Figure 26] Images showing conformational models of wild-type and mutant A305V Enzyme C. Ala305 (green) and Val305 (magenta) are shown in space-filling representation.

[Figure 27] Images showing the surrounding structures of Ala305 of the wild type (left) and Val305 of the mutant (right). Ala305 makes contact with the peripheral amino acid residues. Mutation Val305 causes steric hindrance with the peripheral structure. Ala305 is depicted in green while Val305 is depicted in magenta.

[Figure 28] A diagram showing comparison of intramolecular energies between the wild type (blue) and mutant A305V (red).

[Figure 29] Images showing comparison of the structures between the wild type and mutant A305V. Mutation A305V causes change in the structure of the amino acid residues around Ala305 (green) and Val305 (magenta) (left). This structural change also affects the adjacent  $\alpha$ -helix (right).

[Figure 30] Images showing the effect of A305V in Enzyme C.

[Figure 31] Images showing the effect of A305V in Enzyme C.

[Figure 32] Images showing the effect of A305V in Enzyme C.

[Figure 33] A diagram showing the deduced effect of mutant Enzyme C in the astaxanthin synthesis pathway.

## MODE FOR CARRYING OUT THE INVENTION

[0016] Hereinafter, the present invention will be described in detail.

### 1. General

The present invention relates to a high carotenoid producing bacterium, which is a bacterium comprising either or both of genes (a) and (b) below.

(a) a gene encoding a protein comprising a mutant amino acid sequence in which at least the 225th amino acid residue in the amino acid sequence of 1-deoxy-D-xylulose 5-phosphate synthase of a carotenoidogenic bacterium has been substituted with other amino acid residue; and

(b) a gene encoding a protein comprising a mutant amino acid sequence in which at least the 305th amino acid residue in the amino acid sequence of decaprenyl diphosphate synthase of a carotenoidogenic bacterium has been substituted with other amino acid residue.

[0017] In order to develop a bacterium having high carotenogenic capacity, the present inventors have examined the carotenogenic capacity in strain E-396 and strains thereof that had been subjected to a mutation treatment, and analyzed the mutation of a gene encoding an enzyme involved in the carotenoid synthesis pathway in these strains.

As a result, a strain (referred to as “strain ASB-57”) that had higher carotenogenic capacity than the parent strain E-396 was acquired. Genome analysis for strain ASB-57 confirmed mutations in the amino acid sequence of 1-deoxy-D-xylulose 5-phosphate synthase (DXS) and the amino acid sequence of decaprenyl diphosphate synthase (DPS). Thus, a functional analysis was conducted by prediction based on a conformational analysis of the amino acids, by which the mutation of at least the 225th amino acid residue of DXS and/or the 305th amino acid residue of DPS was considered to contribute to the high carotenoid production.

The present invention was accomplished based on the above-described findings.



[0018] 2. Mutant carotenoidogenic bacterium

A carotenoidogenic bacterium of the present invention is a mutant bacterium that can produce a carotenoid with high efficiency, which can be obtained by subjecting a parent strain to a mutation treatment and using the mutations at the 225th amino acid residue of DXS and/or the 305th amino acid residue of DPS as indicators . A carotenoidogenic bacterium of the present invention is herein referred to as a “mutant carotenoidogenic bacterium”.

(1) Parent strain

According to the present invention, a bacterium used as the parent strain for obtaining a mutant carotenoidogenic bacterium is not limited at all as long as it produces a carotenoid, and may be, for example, a bacterium belonging to the genus *Paracoccus*, the genus *Brevundimonas* or the genus *Erythrobacter*.

A bacterium belonging to the genus *Paracoccus*, a bacterium belonging to the genus *Brevundimonas* or a bacterium belonging to the genus *Erythrobacter* can preferably be used, and a bacterium belonging to the genus *Paracoccus* can more preferably be used. Since all of the genus *Paracoccus*, the genus *Erythrobacter* and the genus *Brevundimonas* belong to the class *Alphaproteobacteria* in the phylum *Proteobacteria*, and are common in bacterial taxonomy, bacteria belonging to these genera can be used for the present invention.

[0019] Among the bacteria belonging to the genus *Paracoccus*, *Paracoccus carotinifaciens*, *Paracoccus marcusii*, *Paracoccus haeundaensis* and *Paracoccus zeaxanthinifaciens* can preferably be used, and *Paracoccus carotinifaciens* can particularly preferably be used. Specific examples of the strains of the bacteria belonging to the genus *Paracoccus* include strain E-396 of *Paracoccus carotinifaciens* (FERM BP-4283) and strain A-581-1 of the bacteria of the genus *Paracoccus* (FERM BP-4671), whose mutants can also preferably be used for the present invention.

Examples of the carotenoidogenic bacteria belonging to the genus *Erythrobacter* include *Erythrobacter* JPCC M sp. (Japanese Patent Application Publication No. 2008-259452) and *Erythrobacter* JPCC O sp. (Japanese Patent Application Publication No. 2008-259449).

Examples of carotenoidogenic bacteria belonging to the genus *Brevundimonas* include *Brevundimonas* sp. strain SD212 (Japanese Patent Application Publication No.

2009-27995), *Brevundimonas* sp. strains FERM P-20515 and 20516 (Japanese Patent Application Publication No. 2006-340676), and *Brevundimonas vesicularis* (Gene, Vol.379, p.101-108, 1 Sep 2006).

[0020] In addition, a bacterium whose nucleotide sequence of DNA corresponding to 16S ribosomal RNA is highly homologous with the nucleotide sequence of strain E-396 represented by SEQ ID NO:9 is preferably used as the carotenoidogenic bacterium. Herein, the homology of the nucleotide sequence is preferably 95% or more, more preferably 96% or more, still more preferably 97% or more, particularly preferably 98% or more, and most preferably 99% or more.

The nucleotide sequence of DNA corresponding to 16S ribosomal RNA refers to a nucleotide sequence obtained by substituting U (uracil) of the nucleotide sequence of 16S ribosomal RNA with T (thymine).

[0021] Classification of microorganisms based on the homology of this nucleotide sequence of 16S ribosomal RNA has been the recent mainstream. Since conventional classification of microorganisms is based on conventional microbiological characteristics such as motility, auxotrophy, sugar assimilation and the like, a microorganism may be classified incorrectly when morphological change or the like is caused due to naturally occurring mutation. On the other hand, the nucleotide sequence of 16S ribosomal RNA is genetically very stable and thus reliability of the classification technique based on that homology is considerably enhanced compared to that of the conventional classification technique.

[0022] Homologies of the nucleotide sequence of 16S ribosomal RNA of *Paracoccus carotinifaciens* strain E-396 with the nucleotide sequences of 16S ribosomal RNA of other carotenoidogenic bacteria, i.e., *Paracoccus marcusii* strain DSM 11574, strain N-81106 of the bacterium of the genus *Paracoccus*, *Paracoccus haeundaensis* strain BC 74171, strain A-581-1 of the bacterium of the genus *Paracoccus*, *Paracoccus zeaxanthinifaciens* strain ATCC 21588, and *Paracoccus* sp. strain PC-1 are 99.7%, 99.7%, 99.6%, 99.4%, 95.7% and 95.4%, respectively, which represent that these strains are closely related taxonomically. Hence, these strains are found to form a group of bacteria that produce carotenoids. Therefore, these strains can preferably be used for the present invention, and are capable of efficiently producing carotenoids.

[0023] According to the present invention, known mutants having improved carotenoid productivity can also be used. Examples of such known mutants include a strain having high astaxanthin production capacity (Japanese Patent Application Publication No. 2001-95500), a strain that selectively produce an increased amount of canthaxanthin (Japanese Patent Application Publication No. 2003-304875), a strain that selectively produce increased amounts of zeaxanthin and  $\beta$ -cryptoxanthin (Japanese Patent Application Publication No. 2005-87097), and a strain that selectively produce an increased amount of lycopene (Japanese Patent Application Publication No. 2005-87100).

[0024] Strain E-396 as an exemplarily carotenoidogenic bacterium used as the parent strain for the present invention has been internationally deposited to the International Patent Organism Depositary (NITE-IPOD), the National Institute of Technology and Evaluation (NITE) as follows.

International depositary authority: International Patent Organism Depositary,  
National Institute of Technology and Evaluation (NITE)

2-5-8, Kazusakamatari, Kisarazu-shi, Chiba, 292-0818, JAPAN

Identification reference: E-396

Accession number: FERM BP-4283

Date of the original deposit: April 27, 1993

[0025] In addition, strain A-581-1 as another exemplarily carotenoidogenic bacterium used as the parent strain in the present invention has been internationally deposited to the above-mentioned authority as follows.

Identification reference: A-581-1

Accession number: FERM BP-4671

Date of the original deposit: May 20, 1994

[0026] (2) Mutation treatment and screening

A mutant carotenoidogenic bacterium of the present invention can be obtained by subjecting the above-described parent strain to a mutation treatment and using the mutations at the 225th amino acid residue of DXS and/or the 305th amino acid residue of DPS as indicators.

A method of the mutation treatment is not particularly limited as long as it can induce mutation. For example, a chemical method using a mutation agent such as

N-methyl-N'-nitro-N-nitrosoguanidine (NTG) or ethyl methanesulfonate (EMS), a physical method such as ultraviolet irradiation or X-ray irradiation, a biological method using gene recombination or transposon, or the like may be employed. Although the bacterium subjected to the mutation treatment is not particularly limited, it is preferably a carotenoidogenic bacterium.

Moreover, according to the present invention, a protein having the above-described mutation can be prepared by introducing point mutation into a gene (DNA) encoding said protein. For such a mutagenesis method, a mutagenesis kit utilizing a site-directed mutagenesis method such as Kunkel method or Gapped duplex method, for example, QuikChange<sup>TM</sup> Site-Directed Mutagenesis Kit (from Stratagene), GeneTailor<sup>TM</sup> Site-Directed Mutagenesis System (from Invitrogen), TaKaRa Site-Directed Mutagenesis System (Mutan-K, Mutan-Super Express Km, etc: from Takara Bio) or the like can be used. In addition, a method such as site-directed mutagenesis described in "Molecular Cloning, A Laboratory Manual (4th edition)" (Cold Spring Harbor Laboratory Press (2012)) or the like can be used.

[0027] While a method for screening a mutant is not particularly limited, a known genome analysis tool such as PacBio RS II (from Pacific Biosciences) or MiSeq (from Illumina) can be used for gene analysis to confirm the presence of a mutation of the nucleotide sequence corresponding to the 225th amino acid residue of DXS and/or the 305th amino acid residue of DPS.

Furthermore, along with the above-described genome analysis, the mutant of interest can, for example, be selected by the color tone of the colony on the agar medium, or selected by culturing the mutants in a test tube, a flask, a fermentation tank or the like and using the production amount of the carotenoid as an indicator in an analysis of carotenoid pigment utilizing absorbance, high performance liquid chromatography, thin-layer chromatography or the like.

The mutation and screening steps can be carried out once, or the mutation and screening steps can be repeated twice or more, for example, by obtaining mutants by a mutation treatment and screening and subjecting the resultant to another mutation treatment and screening to acquire a mutant with improved productivity.

[0028] The mutant carotenoidogenic bacterium screened as such has a gene encoding an

amino acid sequence in which the 225th amino acid residue of DXS is altered to other amino acid and/or the 305th amino acid residue of DPS is altered to other amino acid residue.

Mutation of the 225th amino acid residue of DXS to other amino acid contributes to the increase in the enzymatic activity of DXS. Thus, synthesis from pyruvic acid to 1-deoxy-D xylulose-5-phosphate is promoted, which leads to the increase in the production of isopentenyl diphosphate (IPP) that serves as a substrate for astaxanthin synthesis.

Mutation of the 305th amino acid residue of DPS to other amino acid residue contributes to the reduction in the enzymatic activity of DPS. This mutation suppresses synthesis from farnesyl diphosphate (FPP) to decaprenyl diphosphate (DPP). Since IPP is used for the synthesis of DPP from FPP, the above-described mutation will reduce the amount of IPP used for the DPP synthesis and said IPP will be utilized as a substrate for the above-described astaxanthin synthesis.

[0029] Here, according to the present invention, as long as an amino acid sequence has the 225th amino acid residue of DXS altered to other amino acid and/or the 305th amino acid residue of DPS altered to other amino acid residue, and has a gene coding for a protein comprising an amino acid sequence having such DXS activity and/or a protein comprising an amino acid sequence with reduced (suppressed) DPS activity, one or more amino acid residues in an amino acid sequence of other region of the amino acid sequence of DXS and/or DPS may be altered by substitution, deletion, addition or the like.

Accordingly, a mutant carotenoidogenic bacterium of the present invention may comprise the gene (a) below, the gene (b) below, or both of the genes (a) and (b) below.

[0030] (a) A gene encoding a protein comprising a mutant amino acid sequence in which at least the 225th amino acid residue has been substituted with other amino acid residue in the amino acid sequence of DXS of a carotenoidogenic bacterium.

Examples of such mutant DXS genes include the followings.

(i) A gene encoding a protein that comprises a mutant amino acid sequence in which the 225th amino acid residue in the amino acid sequence of DXS (for example, SEQ ID NO:2) has been substituted with other amino acid residue, and that has DXS activity.

An example of such a mutant amino acid sequence includes one represented by SEQ ID NO:6, while an example of the above-mentioned gene includes one represented by SEQ ID NO:5. According to the present invention, it is preferably an amino acid sequence

in which glycine as the 225th amino acid residue in the amino acid sequence represented by SEQ ID NO:2 is substituted with aspartic acid.

(ii) A gene encoding a protein that comprises a mutant amino acid sequence in which the 225th amino acid residue in the amino acid sequence of DXS (for example, SEQ ID NO:2) is substituted with other amino acid residue and in which one or more (for example, one to several) amino acid residues other than the 225th amino acid residue have been deleted, substituted or added, and that has DXS activity.

(iii) A gene consisting of DNA comprising the nucleotide sequence represented by SEQ ID NO:5.

(iv) A gene comprising DNA that hybridizes with DNA consisting of a nucleotide sequence complementary to DNA comprising the nucleotide sequence represented by SEQ ID NO:5 under stringent conditions, and that encodes a protein having DXS activity.

Among the DNA encoding the amino acid sequence of DXS (SEQ ID NO:1) of a carotenoidogenic bacterium, the above-described nucleotide sequence represented by SEQ ID NO:5 is one that codes for a protein comprising an amino acid sequence in which the 225th amino acid residue has been substituted with other amino acid residue.

[0031] (b) A gene encoding a protein comprising a mutant amino acid sequence in which at least the 305th amino acid residue in the amino acid sequence of DPS of a carotenoidogenic bacterium has been substituted with other amino acid residue.

Examples of such genes include the followings.

(i) A gene encoding a protein that comprises a mutant amino acid sequence in which the 305th amino acid residue in the amino acid sequence of DPS (for example, SEQ ID NO:4) has been substituted with other amino acid residue, and that has reduced DPS activity.

An example of such a mutant amino acid sequence includes one represented by SEQ ID NO:8, while an example of the above-described gene includes one represented by SEQ ID NO:7. According to the present invention, it is preferably an amino acid sequence in which alanine as the 305th amino acid residue in the amino acid sequence represented by SEQ ID NO:4 is substituted with valine.

(ii) A gene encoding a protein that comprises a mutant amino acid sequence in which the 305th amino acid residue in the amino acid sequence of DPS (for example, SEQ

ID NO:4) is substituted with other amino acid residue and in which one or more (for example, one to several) amino acid residues other than the 305th amino acid residue have been deleted, substituted or added, and that has reduced DPC activity.

(iii) A gene consisting of DNA comprising the nucleotide sequence represented by SEQ ID NO:7.

(iv) A gene comprising DNA that hybridizes with DNA consisting of a nucleotide sequence complementary to DNA comprising the nucleotide sequence represented by SEQ ID NO:7 under stringent conditions, and that encodes a protein having reduced DPS activity.

Among the DNA encoding the amino acid sequence of DPS (SEQ ID NO:3) of a carotenoidogenic bacterium, the above-described nucleotide sequence represented by SEQ ID NO:7 is one that codes for a protein comprising an amino acid sequence in which the 305th amino acid residue has been substituted with other amino acid residue.

[0032] Here, the hybridization can be performed according to a known method (for example, Sambrook J. et al., Molecular Cloning, A Laboratory Manual (4th edition) (Cold Spring Harbor Laboratory Press (2012))). Highly stringent conditions refer to conditions under which so-called specific hybrids, but not non-specific hybrids are formed, which are, for example, a sodium concentration of 10 mM-300 mM, preferably 20 mM-100 mM and a temperature of 25°C-70°C, preferably 42°C-55°C.

[0033] Examples of such a mutant carotenoidogenic bacterium include strain ASB-57, strain ASK-8 and strain ASH-66.

Strain ASB-57 has a gene encoding a protein comprising an amino acid sequence in which glycine as the 225th amino acid residue of DXS is altered to aspartic acid, and alanine as the 305th amino acid residue of DPS is altered to valine. The amino acid sequence of DXS in strain ASB-57 and the nucleotide sequence of the gene thereof are represented by SEQ ID NOS:6 and 5, respectively. Moreover, the amino acid sequence of DPS in strain ASB-57 and the nucleotide sequence of the gene thereof are represented by SEQ ID NOS:8 and 7, respectively

[0034] (4) Preparation of gene recombinant

According to the present invention, the gene encoding the above-described mutant DXS and/or the gene encoding the above-described mutant DPS is introduced into a host for transformation, thereby obtaining a gene recombinant type mutant carotenoidogenic

bacterium.

A recombinant vector can be obtained by introducing the mutant DXS gene and/or the mutant DPS gene into a vector, and a transformant can be obtained by introducing said recombinant vector into a host by employing any known method such as Sambrook J. et al., Molecular Cloning, A Laboratory Manual (4th edition) (Cold Spring Harbor Laboratory Press (2012)).

When synthesizing the above-described DXS gene and DPS gene by genetic engineering, DNA coding for said enzyme is first designed and synthesized. The design and synthesis of DNA can be conducted, for example, by a PCR method using a vector containing the full-length gene or the like as a template and primers designed to synthesize the DNA region desired. Then, the above-described DNA is linked with a suitable vector to obtain a recombinant vector for protein expression, and this recombinant vector is introduced into a host such that the gene of interest is expressed, thereby obtaining a transformant (Sambrook J. et al., Molecular Cloning, A Laboratory Manual (4th edition) (Cold Spring Harbor Laboratory Press (2012))).

As the vector, a phage or a plasmid that can autonomously grow in a host microorganism is used. Alternatively, an animal virus or an insect virus vector can also be used. A recombinant vector can be prepared by cleaving purified DNA with a suitable restriction enzyme, and inserting the resultant at a restriction enzyme site of suitable vector DNA or the like to link with the vector. The host used for transformation is not particularly limited as long as it is capable of expressing the gene of interest. For example, it may be a bacterium (*Bacillus subtilis*, bacteria of the genus *Paracoccus*, etc.), a yeast, animal cells (COS cells, CHO cells, etc.), plant cells, insect cells or an insect. The method for introducing a recombinant vector into a host is known.

Moreover, the method for introducing a mutation into gene is the same as described above.

### [0035] (3) Production of carotenoid

According to the present invention, the above-described carotenoidogenic bacterium or transformant can be cultured in a predetermined medium to stably produce a carotenoid at a high concentration.

While the produced carotenoid is not particularly limited, it may be, for example,



astaxanthin, canthaxanthin, zeaxanthin,  $\beta$ -cryptoxanthin, lycopene,  $\beta$ -carotene, adonirubin, adonixanthin, echinenone, asteroidenone or 3-hydroxyechinenone, preferably astaxanthin, canthaxanthin, zeaxanthin or  $\beta$ -cryptoxanthin, and more preferably astaxanthin or zeaxanthin. The carotenoids produced by the present invention may be of a single type or a combination of multiple types.

[0036] Hereinafter, a method for culturing the mutant carotenoidogenic bacterium or transformant of the present invention will be described.

[0037] A carotenoid-producing medium used for culture of the present invention may be added with any component as long as it allows growth of a carotenoidogenic bacterium or transformant and production of a carotenoid. Although any medium containing such an additive can be used, it is preferably a medium containing a carbon source, a nitrogen source, an inorganic salt and if necessary a vitamin or the like.

[0038] Examples of the carbon source include sugars such as glucose, sucrose, lactose, fructose, trehalose, mannose, mannitol and maltose, organic acids such as acetic acid, fumaric acid, citric acid, propionic acid, malic acid, malonic acid and pyruvic acid, alcohols such as ethanol, propanol, butanol, pentanol, hexanol, isobutanol and glycerol, fats and oils such as soybean oil, rice bran oil, olive oil, corn oil, sesame oil and linseed oil. Among them, glucose or sucrose is preferably used. Among these carbon sources, one or more types can be used. While the amount added to the medium before culture (starting medium) varies according to the type of the carbon source and can appropriately be adjusted, it is usually 1-100 g, preferably 2-50 g per 1L of the medium. Furthermore, the carbon source may be added not only to the starting medium but also preferably additionally supplied during culture successively or continuously.

[0039] As an inorganic salt as the nitrogen source, one or more types among ammonium salts such as ammonium nitrate, ammonium sulfate, ammonium chloride and ammonium phosphate, nitrates such as potassium nitrate, ammonia and urea can be used. While the amount added varies and can appropriately be adjusted according to the type of the nitrogen source, it is usually 0.1 g-20 g and preferably 0.2-10 g per 1L of the medium.

Furthermore, as an organic nitrogen source, for example, one or more types among corn steep liquor (including filtrated product), Pharmamedia, soybean pulp, soybean powder, peanut meal, soy peptone, Distillers' solubles, dry yeast, yeast extract, casamino acid,

glutamic acid and aspartic acid can be used. While the concentration added varies and can appropriately be adjusted according to the type of the nitrogen source, it is usually 0-80 g/L and preferably 1-30 g/L.

[0040] The inorganic nitrogen source and the organic nitrogen source are usually added to the starting medium, but they may also preferably be additionally supplied successively or continuously.

As the inorganic salt, for example, one or more types among phosphates such as potassium dihydrogen phosphate, dipotassium hydrogen phosphate and disodium hydrogen phosphate, magnesium salts such as magnesium sulfate and magnesium chloride, iron salts such as iron sulfate and iron chloride, calcium salts such as calcium chloride and calcium carbonate, sodium salts such as sodium carbonate and sodium chloride, manganese salts such as manganese sulfate, copper salts such as copper sulfate, zinc salts such as zinc sulfate, molybdenum salts such as sodium molybdate, nickel salts such as nickel sulfate, selenium salts such as sodium selenate, tungsten salts such as sodium tungstate, aluminum salts such as aluminum chloride, chromium salts such as chromium chloride, and potassium borate and iodide can be used. While the amount added varies and can appropriately be adjusted according to the type of the inorganic salt, it is usually 0.0001-15 g per 1L of the medium. The concentration is preferably 0.02-15 g/L in a case of a phosphate, a magnesium salt, a calcium salt, a sodium salt or an iron salt, and preferably 0.1-15 mg/L when a manganese salt, a copper salt, a zinc salt, a molybdenum salt, a nickel salt, a selenium salt, a tungsten salt, an aluminum salt, a chromium salt, or a potassium borate or iodide is added. An inorganic salt is usually added to the starting medium, but it may also be additionally supplied successively or continuously.

[0041] As a vitamin, for example, cyanocobalamin, riboflavin, pantothenic acid, pyridoxine, thiamine, ascorbic acid, folic acid, niacin, p-aminobenzoic acid, biotin, inositol, choline or the like can be used. While the amount added varies and can appropriately be adjusted according to the type of the vitamin, it is usually 0.001-1000 mg and preferably 0.01-100 mg per 1L of the medium. A vitamin is usually added to the starting medium, but it may also be additionally supplied successively or continuously.

[0042] According to the present invention, an antifoaming agent is preferably used in order to suppress foaming of the culture solution. Any type of antifoaming agent can be

used as long as it serves to suppress generation of foam or eliminate generated foam, and has little inhibitory action against the produced bacterium. For example, an alcohol-based antifoaming agent, a polyether-based antifoaming agent, an ester-based antifoaming agent, a fatty acid-based antifoaming agent, a silicon-based antifoaming agent, a sulfonic acid-based antifoaming agent, and the like can be exemplified. While the amount added varies and can appropriately be adjusted according to the type of the antifoaming agent, it is usually 0.01 g-10 g per 1L of the medium.

[0043] The antifoaming agent is usually added to the starting medium prior to sterilization. In addition, it may continuously or intermittently be added during the culture. As a method of adding an antifoaming agent during the culture, a method in which the antifoaming agent is automatically added once a sensor senses foam generation, a method in which the antifoaming agent is added constantly using a programmable timer, a method in which the antifoaming agent is added as a mixture with a feeding carbon source, nitrogen source or pH regulator or the like so as to link with the growth rate, or the like can be exemplified. An antifoaming agent added to the starting medium and an antifoaming agent added during the culture may be of the same type or they may be different according to their actions.

[0044] According to the present invention, the early pH of the medium is adjusted to 2-12, preferably 6-9 and more preferably 6.5-8.0. The pH in the above-mentioned range is preferably maintained during the culture as well. As the pH regulator, an aqueous sodium hydroxide solution, an aqueous potassium hydroxide solution, an aqueous sodium carbonate solution, ammonia water, ammonia gas, an aqueous sulfate solution or a mixture thereof can be exemplified.

According to the present invention, the medium is subjected to a sterilization treatment and thereafter used for culturing a bacterium. The sterilization treatment can appropriately be carried out by those skilled in the art. For example, the medium in a suitable vessel may be heat sterilized with an autoclave. Alternatively, it may be sterilized by filtration with a sterilizing filter.

[0045] A mutant carotenoidogenic bacterium or transformant of the present invention is inoculated onto a medium prepared as described above and cultured under predetermined conditions. Inoculation is carried out by appropriately growing the bacterial strain by seed

culture using a test tube, a flask, a fermentation tank or the like, and adding the resulting culture solution to a carotenoid-producing medium. The medium used for the seed culture is not particularly limited as long as it is a medium that allows the carotenoidogenic bacterium to grow well.

[0046] Culture is carried out in a suitable culture vessel. While the culture vessel may appropriately be selected according to the culture volume, it may be, for example, a test tube, a flask and a fermentation tank.

The culture temperature is 15-40°C, preferably 20-35°C and more preferably 25°C-32°C and culture is usually conducted for 1-18 days, preferably 2-12 days and more preferably 3-8 days under an aerobic condition. The aerobic condition, for example, may refer to shaking culture, aeration-agitation culture or the like, where the dissolved oxygen concentration is preferably controlled to lie within a certain range. The dissolved oxygen concentration can be controlled, for example, by varying the rotation speed for agitation, the ventilation volume, the internal pressure or the like. The dissolved oxygen concentration is preferably controlled to 0.3-10 ppm, more preferably 0.5-7 ppm and still more preferably 1-5 ppm.

[0047] The number of bacterial cells or the number of transformants of a carotenoidogenic bacterium after culturing the mutant carotenoidogenic bacterium or transformant of the present invention can be measured by OD. In addition, the carotenoid contained in the resulting cultured product or the carotenoid collected from the cultured product after culturing the carotenoidogenic bacterium or transformant can be quantified by high performance liquid chromatography. After culturing the carotenoidogenic bacterium or transformant, the carotenoid can be collected from the resulting cultured product as described above.

Examples of the cultured product include a culture solution, a culture supernatant, a concentrated bacterial cell solution, wet bacterial cells, dry bacterial cells and a bacterial cell lysate. The culture supernatant can be prepared by subjecting the culture solution to a centrifugal treatment or a filtration treatment to remove the bacterial cells from the culture solution. The concentrated bacterial cell solution can be obtained by subjecting the culture solution to centrifugation or membrane filtration concentration. The wet bacterial cells can be obtained by centrifuging or filtrating the culture solution. The dry bacterial cells can be

obtained by drying wet bacterial cells or a concentrated bacterial cell solution by a general drying method. The resulting carotenoid-containing dry bacterial cells can directly be used as feed additives.

[0048] The yield by fermentation culture is at least 150 mg/L, for example, 150 mg/L, 400 mg/L, 2000 mg/L or 4000 mg/L of carotenoid. While the amount of the carotenoid contained in the culture solution varies depending on the bacterial cells used, it may contain, for example, 400 mg/L-4000 mg/L and still more preferably 500 mg/L-3500 mg/L of carotenoid.

The bacterium of the present invention has production capacity of an amount at least 5 times, preferably 10 times or more the carotenoid production amount of a carotenoidogenic bacterium that does not have the gene encoding the protein comprising the mutant amino acid sequence of DXS and/or DPS.

[0049] According to the present invention, the method for collecting the carotenoid from the above-described cultured product is not particularly limited, and any method that allows stable and efficient carotenoid collection can be employed. These methods can appropriately be selected from known extraction and purification techniques and carried out by those skilled in the art. Alternatively, according to the present invention, the above-described cultured product can also be used as a carotenoid-containing composition.

Prior to extraction, the cultured product may be subjected to one or more treatments among a chemical treatment using an alkali reagent or a surfactant, a biochemical treatment using a lytic enzyme, a lipolytic enzyme or a proteolytic enzyme, or a physical treatment like ultrasonic waves or pulverization.

For example, in a case where a carotenoid is extracted from the cultured product, a solvent used for extraction and washing is not particularly limited, and it may be a lower alcohol such as methanol, ethanol and isopropanol, acetone, tetrahydrofuran, methyl ethyl ketone, methyl isobutyl ketone, dichloromethane, chloroform, dimethyl formamide or dimethyl sulfoxide.

[0050] If oxidization of the carotenoid needs to be minimized during the extraction operation, it may be treated in an inert gas atmosphere such as nitrogen gas. Alternatively, an antioxidant used for pharmaceutical products and food can be selected and added to the extraction solvent. Alternatively, these treatments can be combined. Moreover, in order to

minimize photolysis of the carotenoid, the treatment can be conducted under a condition without exposure to light.

The thus-resulting extract can directly be used as the carotenoid, or it may be purified before use.

[0051] While a method for separating the bacteria or the like remaining in the extract after the extraction operation is not particularly limited, membrane filtration, centrifugation, decantation or the like can be employed.

A method for obtaining the carotenoid precipitate from the extracted liquid generally includes heating and/or concentration under reduced pressure, or crystallization. Alternatively, the carotenoid pigments can be separated without concentration by precipitation of carotenoid pigments at a low temperature, or precipitation with an acidic/alkaline agent or a salt. For an industrial use, crystallization is favorable.

[0052] The resulting carotenoid precipitate may be suspended and agitated using a solvent such as a small amount of a lower alcohol for washing if necessary. While the washing procedure is not particularly limited, examples of a practically preferable method include a method employing filtration following the suspension/agitation, and a method in which a liquid is passed from top of the precipitate.

The resulting cultured product, extract or purified product may be used as a carotenoid alone or as a mixture at arbitrary proportions.

[0053] 2. Conformational analysis of enzyme involved in astaxanthin synthesis pathway

Since the mutations of the 225th amino acid residue of DXS and the 305th amino acid residue of DPS are found to play an important role in carotenoid synthesis, conformational analyses of these enzymes can be conducted.

According to the present invention, point mutations in two types of enzymes (referred to as Enzymes A and C) on the astaxanthin synthesis pathway were identified. Since the increase in the astaxanthin production was considered to result from the mutations caused in these enzymes, conformational models were constructed for the two types of enzymes having the identified mutations to predict the effects of the amino acid substitutions due to the mutations.

[0054] Enzyme A was deduced to be 1-deoxy-D-xylulose 5-phosphate synthase (DXS). The identified amino acid mutation G225D was in the disordered region near the active site.

From the conformational model, mutation G225D in Enzyme A was deduced to cause a structural change in Enzyme A that was similar to that caused by the mutation in the disordered region that was known to enhance DXS activity, thereby predicting enhanced enzymatic activity in the mutant G225D enzyme. Considering that DXS is under the control of feedback inhibition of IPP, the amount of IPP produced is suggested to increase because feedback inhibition does not work in the mutant G225D enzyme. The increase in the supply of IPP, i.e., a raw material of astaxanthin, due to mutation G225D in Enzyme A seems to cause the increase in the astaxanthin production amount.

[0055] Enzyme C was deduced to be decaprenyl diphosphate synthase. From the conformational model, the identified mutation A305V was deduced to cause steric hindrance with the surrounding amino acid residues and destabilizes the conformation of Enzyme C. FPP and IPP as the raw materials of decaprenyl diphosphate as well as the substrates of Enzyme C are also raw materials for the astaxanthin synthesis. A decrease in the Enzyme C activity due to destabilization is suggested to decrease the amounts of FPP and IPP digested by Enzyme C. As a result, the amounts of FPP and IPP that can be used for the astaxanthin synthesis are increased, which is considered to increase the astaxanthin production amount.

[0056] Accordingly, two effects, namely, the increase in the IPP production amount due to the enhancement of the Enzyme A activity and the increase in the IPP supply to the astaxanthin synthesis pathway due to the decrease in the Enzyme C activity, were deduced to be the effects of the mutations identified this time on the astaxanthin synthesis pathway. These effects appear to result in the increase in the astaxanthin production amount.

#### [0057] EXAMPLES

Hereinafter, the present invention will be described more specifically by way of examples. The scope of the present invention, however, should not be limited to these examples.

##### [0058] Example 1

###### (1) Mutation treatment and genome analysis of bacteria of genus *Paracoccus*

###### Method of mutation treatment

Screening was carried out for several times by using UV, NTG (nitrosoguanidine) and the like as mutagens on the parent strain (strain E-396) and adopting various selection

pressures. Screening was conducted using the astaxanthin yield as an indicator.

#### Method of genome analysis

Following genome sequencing using PacBio *RS II* (from Pacific Biosciences) or MiSeq (Illumina) sequencer, a genome analysis was carried out by using an analysis software such as SMART Cell 8 Pac V3 (from Pacific Biosciences), MiSeq Control Software (MCS) v2.4.1.3, Real Time Analysis (RTA) v1.18.54, or bcl2fastq v 1.8.4 (Illumina).

#### [0059] Results from genome analysis (identification of mutation site)

The mutation site was identified by sorting out those belonging to the genus *Paracoccus* among the amino acid sequences of enzyme genes recorded in the Kyoto Encyclopedia of Genes and Genomes (KEGG) which have high homology with the amino acid sequence of the region that was considered to be the protein having the mutation point from the genome analysis, and further conducting a conformational analysis of the enzyme based on these information to find out a template having a common sequence, thereby determining the final enzyme name of the amino acid sequence of the protein having the mutation site.

#### [0060] (2) Astaxanthin yield

##### (i) Culture conditions

Strokes of test tubes: 330 rpm, 28°C, pH 7.2, amount of medium 8 ml/tube

Culture time: 72 hours

Medium:

8 ml of a medium with the following composition was placed in a test tube provided with a cotton plug and having an inner diameter of 18 mm and sterilized in an autoclave at 121°C for 15 minutes to prepare a test tube medium for seeding. The raw materials used for the test tube medium for seeding were those from lots that were confirmed to allow sufficient growth of bacterial cells.

Sucrose: 30 g/L

Corn steep liquor: 30 g/L

Potassium dihydrogen phosphate: 1.5 g/L

Disodium hydrogen phosphate dodecahydrate: 3.8 g/L

Calcium chloride dihydrate: 5.0 g/L



Magnesium sulfate heptahydrate: 0.7 g/L

Iron sulfate heptahydrate: 1.0 g/L

pH7.2

[0061] Next, 7.2 ml of a medium with the following composition was placed into each of test tubes provided with a cotton plug and having an inner diameter of 18 mm to prepare five test tube media for production. The raw materials used for the test tube media for production were those from lots that were confirmed to allow insufficient growth of bacterial cells.

Glucose: 30 g/L

Filtrated corn steep liquor product: 30 g/L

Ammonium sulfate: 1.5 g/L

Potassium dihydrogen phosphate: 1.5 g/L

Disodium hydrogen phosphate dodecahydrate: 3.8 g/L

Calcium chloride dihydrate: 5.0 g/L

Magnesium sulfate heptahydrate: 0.7 g/L

Iron sulfate heptahydrate: 1.0 g/L

Silicon-based antifoaming agent: 0.2 g/L

[0062] (ii) Results

The specific productivity of each strain is shown in Figure 1.

According to this example, strain ASB-57 that had production capacity 10 times or more that of strain E-396 carotenoid was acquired.

[0063] Example 2

Conformational analysis of enzyme involved in astaxanthin synthesis pathway

1. Conformational data and procedure

The conformational models of Enzymes A and C were constructed by homology modeling. For the modeling, software Swiss-Pdb viewer and SWISS-MODEL were used [1, 2]. Mutant models were prepared with Swiss-Pdb viewer. The command “mutate” was used to substitute the amino acid residues, the command “compute energy” was used to compute intramolecular energy, and the command “energy minimization” was used to calculate for energy minimization. Preparation of a complex model with the substrate or the like, detection of the residues in the vicinity of the substrate, measurement of interatomic distance, and display of the conformation were conducted using software Waals (Altif Labs. Inc.). A

conformational model of a low-molecular compound was prepared with MarvinSketch (ChemAxon Ltd.).

[0064] The coordinate data of the conformation of the template structure was acquired from the protein conformational database, Protein DataBase (PDB) (<http://www.rcsb.org/pdb/>). The template structures that had the highest amino acid matching degree with each of Enzymes A and C were used among the data registered with PDB. The conformational data used as the templates for the homology modeling is shown in Table 1.

Table 1: Conformational data used for homology modeling

Enzyme	PDB ID	Protein	Document
Enzyme A	2O1X	1-Deoxy-D-xylulose-5-phosphate synthase	[3]
Enzyme C	3MZV	Decaprenyl diphosphate synthase	[4]

## [0065] 2. Construction of conformational model of Enzyme A and analysis of mutant

Enzyme A is 1-deoxy-D-xylulose 5-phosphate synthase (DXS) that synthesizes 1-deoxy-D xylulose 5-phosphate from pyruvic acid and D-glyceraldehyde 3-phosphate in the deoxyxylulose pathway, i.e., one of isoprenoid biosynthesis pathways that biosynthesize isoprenyl diphosphate (IPP) that serves as a raw material for astaxanthin synthesis.

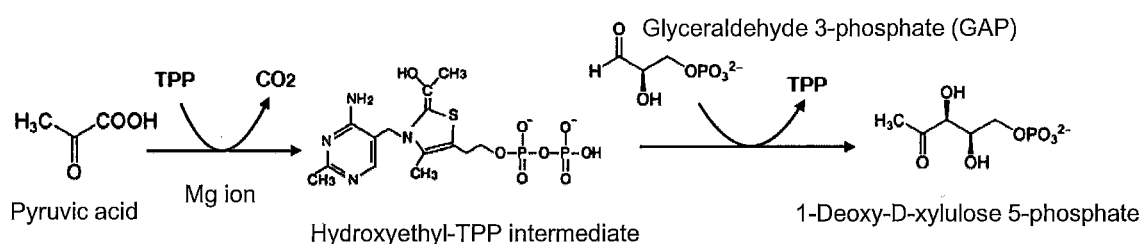
[0066] Mutation G225D identified from the constructed conformational model of Enzyme A by the genome analysis was found to exist in the disordered region in the vicinity of the active site, where the mutation had been reported to enhance the enzymatic activity of DXS in several cases. From the results of the analysis of the mutant model, mutation G225D in Enzyme A was deduced to induce a structural change similar to that caused by the mutation known to enhance the DXS activity, and thus the mutant G225D enzyme was also predicted to have enhanced DXS enzymatic activity similar to the known mutation. Enhancement of the DXS activity increases IPP supply as the raw material of astaxanthin, which is considered to increase the amount of astaxanthin produced.

### [0067] 2.1. Enzymatic reaction of 1-deoxy-D-xylulose 5-phosphate synthase

Enzyme A, 1-deoxy-D-xylulose-5-phosphate synthase (DXS), synthesizes 1-deoxy-D-xylulose 5-phosphate from pyruvic acid and D-glyceraldehyde 3-phosphate in

the presence of a magnesium ion (Mg). The catalytic reaction requires thiamine pyrophosphate (TPP) as a coenzyme. First, TPP coenzyme attaches to the substrate pyruvic acid to give a hydroxyethyl-TPP intermediate. Reaction between this intermediate and glyceraldehyde 3-phosphate generates 1-deoxy-D-xylulose 5-phosphate. The enzymatic reaction of Enzyme A is shown below.

#### Chemical formula 1



### [0068] 2.2. Construction of conformational model of Enzyme A

#### (1) Construction of conformational model of Enzyme A by homology modeling

The conformational model of Enzyme A was constructed by homology modeling based on the conformation (PDB ID:2O1X) [3] of *Deinococcus radiodurans* (*D. radiodurans*)-derived 1-deoxy-D-xylulose-5-phosphate synthase (DXS, template) (Figure 2) whose conformation of the complex with TPP coenzyme had been determined by X-ray crystallography.

[0069] Homology modeling was carried out based on the conformation alignment between Enzyme A and the template structure (Figures 3A and 3B).

The amino acid matching degree between Enzyme A and *D. radiodurans*-derived DXS was 44.1%. In the conformation of the template DXS, the region of the amino acid residues 199-242 (44 residues) was a disordered region and thus the positions of the atoms were unspecified by X-ray crystallography. Therefore, a conformational model of the amino acid residues 7-630 excluding the residues 196-238 (43 residues) corresponding to the disordered region of Enzyme A was constructed. Then, TPP and Mg were embedded by superimposing the conformational model of Enzyme A and the template structure to prepare a complex model of Enzyme A and TPP. Figure 4 shows the template structure and the constructed model structure.

Similar to the template structure, Enzyme A formed a homodimer and had a TPP binding site and a substrate binding site in each of the subunits. The monomer of Enzyme A consists of three domains, namely, domain I (residues 1-319), domain II (residues 320-495) and domain III (residues 496-629).

[0070] (2) Preparation of Enzyme A-substrate complex models

In order to deduce the amino acid residue responsible for the bond between Enzyme A and the substrates, complex models having the substrates bonded to Enzyme A were constructed. Since the coordinates of pyruvic acid and glyceraldehyde 3-phosphate as the substrates were not determined in the template structure 2O1X, first, a model of a complex with the hydroxyethyl-TPP intermediate that had TPP coenzyme attached to pyruvic acid was prepared by embedding the hydroxyethyl-TPP intermediate by superimposition based on a conformation in which a related *Saccharomyces cerevisiae* (*S. cerevisiae*)-derived transketolase (TK) was bound with the hydroxyethyl-TPP intermediate (PDB ID: 1GPU) [6], to detect the pyruvic acid binding site. Similarly, a model of an Enzyme A-glyceraldehyde 3-phosphate complex was prepared by embedding erythrose-4-phosphate based on a conformation of *S. cerevisiae*-derived TK with erythrose-4-phosphate (PDB ID: 1NGS) [7] and further preparing a glyceraldehyde 3-phosphate model from erythrose-4-phosphate.

Figure 5 shows the model structure of the Enzyme A-substrate complex.

[0071] 2.3. Deduction of active site of Enzyme A

In order to deduce the amino acid residue binding the coenzyme and the substrate, interactions between the TPP intermediate, GAP and Mg were examined in the complex model of Enzyme A and the coenzyme.

[0072] (1) Binding site of TPP coenzyme

Similar to the template structure, TPP was located between domain I and domain II in the complex model of Enzyme A and TPP, where the pyrimidine ring of TPP was bound to domain II and the phosphate group was bound to domain I.

[0073] TPP consists of an aminopyrimidine ring, a thiazoline ring and pyrophosphate (Figure 6). The aminopyrimidine ring binds to closely fit inside the binding pocket of Enzyme A. The side chains of Phe396 and Ile369 are bound via hydrophobic interactions to sandwich the aminopyrimidine ring. Especially, a strong bond via  $\pi$  stacking was deduced

between the side chain of Phe396, the cyclic phenyl group, and the aminopyrimidine ring. The main-chain oxygen of Gly120 and the side-chain oxygen of Glu371 formed hydrogen bonds with the nitrogen atom at position 1 in the aminopyrimidine ring. The hydrogen bond between the nitrogen atom at position 1 in the pyrimidine ring of TPP and Glu residue is important for the catalytic reaction via TPP, and is known to be conserved in an enzyme that uses TPP as a coenzyme. The side chain of Arg399 is predicted to form the binding pocket not only via the hydrophobic bond but also via hydrogen bonds with the side chains of the peripheral amino acid residues Glu371 and Ser122. As to the interactions with the thiazoline ring, hydrophobic bonds with Ile184 and Ile369 were observed. As to the interactions with pyrophosphate, hydrogen bonds with the main chain of Gly152, the side chain of Lys285 and the side chain of His79 were deduced.

[0074] The bond with the TPP coenzyme requires Mg. Mg is located between the two phosphate groups of TPP, and Mg was deduced to bind with the side chains of Asp151 and Asn180 and the main chain of Met182. As to DXS, the sequence of GDGX25-30N is known to be conserved as a TPP binding motif [3]. The amino acid sequence Gly150-Asp151-Gly152-Asn180 including Asp151 and Asn180 that were deduced to form bonds with Mg in Enzyme A matches this motif.

[0075] Thus, the carbons in the aminopyrimidine ring and the thiazoline ring of TPP were deduced to have hydrophobic interactions with hydrophobic residues Ile184, Ile369, Phe396 and Arg399 of Enzyme A. Moreover, two nitrogen atoms in the aminopyrimidine ring were deduced to form hydrogen bonds with Gly120 and Glu371. Besides the hydrogen bonds with His79, Gly152 and Lys285, the pyrophosphate was deduced to form hydrogen bonds with Asp151, Asn180 and Met182 via Mg. Figure 7 shows the amino acid residues that were deduced to be responsible for the bonds with TPP.

[0076] (2) Interaction with pyruvic acid

Pyruvic acid as the substrate reacts with TPP and forms a hydroxyethyl-TPP intermediate. From the complex model of Enzyme A and the hydroxyethyl-TPP intermediate, a hydrophobic interaction with Val77 and a hydrogen bond with His432 were deduced as interactions with the pyruvic acid-derived hydroxyethyl group. These amino acid residues were considered to be involved in the bond with pyruvic acid. Figure 8 shows amino acid residues that were predicted to have interactions with the hydroxyethyl groups.

[0077] (3) Interaction with glyceraldehyde 3-phosphate

As the amino acid residues that interact with the substrate glyceraldehyde 3-phosphate (GAP), hydrogen bonds with His48, Tyr393, Arg421, Asp428 and Arg479 were deduced (Figure 9).

His48 and Asp428 form hydrogen bonds with the aldehyde groups of GAP. Tyr393, Arg421 and Arg479 were deduced to form hydrogen bonds with the phosphate groups of GAP.

[0078] The above-described deduced active sites are shown in Tables 2-4.

Table 2: Amino acid residues predicted to have interaction with hydroxyethyl-TTP intermediate detected in Enzyme A

hydroxyethyl-TTP	Enzyme_A		interaction
Amino-Pyrimidine	Gly120	main-chain	hydrogen bond
	Ile369	side-chain	hydrophobic interaction
	Glu371	side-chain	hydrogen bond
	Phe396	side-chain	hydrophobic interaction
	Arg399	side-chain	hydrophobic interaction
Thiazolium	Ile184	side-chain	hydrophobic interaction
	Ile369	side-chain	hydrophobic interaction
PO4	His79	side-chain	hydrogen bond
	Gly152	main-chain	hydrogen bond
	Lys285	side-chain	hydrogen bond
Mg	Asp151	side-chain	hydrogen bond
	Asn180	side-chain	hydrogen bond
	Met182	main-chain	hydrogen bond
Hydroxyethyl	Val77	side-chain	hydrophobic interaction
	His432	side-chain	hydrogen bond

Table 3: Amino acid residues predicted to have interaction with glyceraldehyde 3-phosphate detected in Enzyme A

GAP	Enzyme_A		interaction
Glyceraldehyde	His48	side-chain	hydrogen bond
	Asp428	side-chain	hydrogen bond
PO4	Tyr393	side-chain	hydrogen bond
	Arg421	side-chain	hydrogen bond
	Arg479	side-chain	hydrogen bond

Table 4: Amino acid residues at deduced binding sites in Enzyme A

binding site	residue
TTP	His79
	Gly120
	Gly152
	Ile184
	Lys285
	Ile369
	Glu371
	Phe396
	Arg399
Mg	Asp151
	Asn180
	Met182
Pyrvate	Val77
	His432
GAP	His48
	Asp428
	Tyr393
	Arg421
	Arg479

[0079] Furthermore, the binding modes between Enzyme A and TPP coenzyme, the binding modes between Enzyme A and pyruvic acid, the binding modes between Enzyme A and glyceraldehyde 3-phosphate are shown in Figures 6, 10 and 11, respectively.

Since Enzyme A retained active sites that bind TPP coenzyme and substrates pyruvic acid and GAP like template *D. radiodurans*-derived DXS, it was predicted to have the enzymatic activity of DXS. Other than *D. radiodurans*-derived DXS, these amino acid residues are also known to be highly conserved in DXS from *E. coli* and the like and *S.*



*cerevisiae*-derived TK [3].

[0080] In *E. coli* DXS, substitution of the amino acid residues corresponding to Glu370, Arg399 and Arg479 of Enzyme A to Ala is found to result deactivation [3]. According to Document [8], enzymatic activity was generally deactivated in all of the experiments of mutating amino acid residues of *E. coli* DXS corresponding to His48, Glu371 and Asp428 of Enzyme A. These amino acid residues are predicted to be important for DXS activity in Enzyme A as well.

[0081] 2.4. Effect of mutation G225D in Enzyme A

Mutation G225D in Enzyme A was present in the disordered region (residues 196-238) whose conformation was unspecified. In order to deduce the effect of this mutation on the conformation of Enzyme A, the previous findings, the conformational location of the disordered region and the relationship with the active site were examined.

[0082] (1) Regarding mutation in known disordered region

So far, two mutations caused in the disordered region of each of Muscat and *E. coli* DXS, i.e., a total of four mutations, were all reported to increase the enzymatic activity.

In Muscat (*Vitis vinifera*)-derived DXS, mutation K284N was reported to result an increase in the activity that was about twice as high as that of the wild-type in Vmax and Kcat/Km, and overexpression was reported to greatly increase the amount of monoterpene produced [9].

[0083] Moreover, inventions related to mutations K284N and R306C in Muscat and mutations K213N and K234C in *E. coli* are known (Japanese Patent Application Publication No. 2014-500710, US20130276166). This invention relates to a method for increasing the amount of terpene produced by enhancement of DXS activity, where the amount of terpene produced increased by the single-residue mutation in all of the four cases.

[0084] The positions of mutations in Enzyme A and *E. coli* and Muscat DXS that resulted increase in the enzymatic activity are shown in Figure 12.

K284N and R306C in Muscat, and K213N and K234C in *E. coli* all existed in the disordered region (blue). Mutation G225D in Enzyme A also existed in the disordered region. Based on the amino acid sequence alignment, the active site (green) of DXS was conserved and Enzyme A was predicted to have a reaction mode similar to the reaction modes of these DXS. The active site existed on the N-terminal side (magenta) of the

disordered region.

[0085] Thus, multiple mutations in the disordered region of DXS increased the enzymatic activity of DXS, and mutagenesis in this region of interest was predicted to give some effect to the DXS activity.

[0086] (2) Conformational location of mutation G225D (disordered region) in Enzyme A

Figure 13 shows the location of the disordered region (residues 196-238) that has mutation G225D in Enzyme A in a blue dotted line.

The region of interest is located in the vicinity of the binding site of TPP coenzyme essential for DXS activity. Active sites Asn180, Met182 and Ile184 exist in the loop (magenta) on the N-terminal side of the region of interest. The side chain of Asn180 and the main chain of Met182 bind to Mg. The side chain of Ile184 form a hydrophobic bond with TPP. It is considered important that this loop has a suitable structure such that TPP essential for DXS activity can bind with Mg. In document [9], the physiological role of the region of interest was unclear but the region of interest was considered to exist near the active site, and the mutation in the region of interest rationally seemed to have an effect on the activity of the enzyme. Since the region of interest exists in the vicinity of the TPP binding site both in terms of the amino acid sequence and the conformation as shown in Figures 12 and 13, it is highly likely that the mutation in the region of interest has an effect on these active sites.

[0087] (3) Preparations of model of disordered region of Enzyme A and mutant G225D enzyme

Next, in order to examine the effect of mutation G225D in Enzyme A on the conformation of the disordered region, a model structure was prepared for the disordered region across residues 196-238 (43 residues) of Enzyme A. In document [9] that reported mutation K284N in Muscat, a model of the disordered region of Muscat DXS was prepared to observe the change in the electrostatic potential by the mutation. For reference, a similar analysis was conducted for Enzyme A as well. A conformational model was prepared by homology modeling based on a fragment structure of an amino acid sequence that was highly homologous with the disordered region of Enzyme A. As a template structure, a fragment 1AL7 having the highest amino acid matching degree (34%) among the conformations registered in PDB was used. Even though the disordered region was predicted to have a fluctuated structure, it was considered to serve as a reference of a folding

that is likely to form the region of interest. The prepared model is shown in Figure 14 (left).

[0088] Furthermore, a conformational model of the mutant G225D enzyme was prepared by substituting the mutation site Gly225 with Asp. Since Gly225 is located on the surface, substitution with Asp exposes the side chain of Asp on the surface. The results from mapping electrostatic potential on the surface profile of the prepared conformational model are shown in Figure 14 (right). Blue represents a positive (positively charged) region while red represents a negative (negatively charged) region.

[0089] While there is a strong positively charged region in the wild type, a weakened positive charge and a strengthened negative charge can be confirmed in the mutant G225D enzyme. In the wild type, the side chains of positively charged Arg227, Arg228, Lys230 and K234 aggregate, and form a strong positively charged region. Since the uncharged Gly225 existing in this region was replaced with the negatively charged Asp, the positive charge near Asp225 seemed to have weakened. This result shows similar tendency to document [9], that is, the change in the electrostatic potential due to mutation K284N in Muscat, i.e., the change in the electrostatic potential from a positive charge to a negative charge on the surface of the disordered region. The structural change in the disordered region caused by G225D in Enzyme A was predicted to have an effect similar to the mutation in Muscat on the active site including the TPP binding site, suggesting that the mutant G225D of Enzyme A had an increased enzymatic activity similar to the mutant K284N of Muscat due to this effect.

#### [0090] 2.5. Effect of mutation G225D in Enzyme A on astaxanthin synthesis pathway

Accordingly, mutation G225D in Enzyme A was confirmed to occur in the disordered region of Enzyme A, and this region was confirmed to exist in the vicinity of the TPP binding site that was essential for the activity. So far, multiple mutations in this region have been found to enhance the enzymatic activity of DXS, suggesting that G225D mutation causes a structural change similar to that caused by the mutations that are known to enhance the DXS activity.

[0091] From the location of the disordered region confirmed with the conformational model, mutation in the disordered region seems to have some effect on the TPP binding region to enhance the enzymatic activity at least as described in the document. In the case of Muscat,  $K_{cat}/K_m$  of the mutant was shown to be about twice as high as that of the wild type

by an enzymonological experiment *in vitro*, and thus the mutation in the disordered region was predicted to result a structural change that was more suitable for the binding of TPP.

[0092] In order to greatly increase a monoterpene in a cell, isopentenyl diphosphate (IPP) as a raw material of the monoterpene needs to be increased. Feedback inhibition of DXS by IPP as a product of the deoxyxylulose pathway and competitive inhibition of IPP with TPP have been reported [10]. Moreover, it is known that once IPP reaches a certain amount, DXS, i.e., the first enzyme of the deoxyxylulose pathway, is inhibited so that IPP can no longer increase.

[0093] Since the amount of the IPP supplied is regulated to stay at a certain level by DXS due to this feedback inhibition, an increase in the enzymatic activity  $K_{cat}/K_m$  of DXS would not simply increase the IPP supply to significantly increase terpene. In a known example, the mutation in the disordered region increased the amount of the synthesized terpene and this mutation was predicted to increase the IPP supply, suggesting that feedback inhibition of DXS by IPP was no longer effective.

[0094] IPP and TPP competitively bind to DXS *in vitro*, and are found to inhibit DXS [10]. The mutation in the disordered region that caused a structural change in the TPP binding region suggests not only that it gives a structure suitable for TPP binding but also that feedback inhibition by IPP no longer works because of the effect on IPP binding.

[0095] Mutation G225D in Enzyme A was also predicted to result a structural change of the TPP binding region and increase the amount of IPP produced because feedback inhibition by IPP no longer works. As a result, the supply of IPP as a raw material of astaxanthin synthesis increased, by which the amount of synthesized astaxanthin was increased (Figure 15).

[0096] 3. Construction of conformational model of Enzyme C and analysis of mutant

From the homology of the amino acid sequences and a comparative conformational analysis, Enzyme C was found to be one kind of polyprenyl diphosphate synthases, namely, decaprenyl diphosphate synthase that synthesizes decaprenyl diphosphate from farnesyl diphosphate (FPP) and seven isopentenyl diphosphates (IPP).

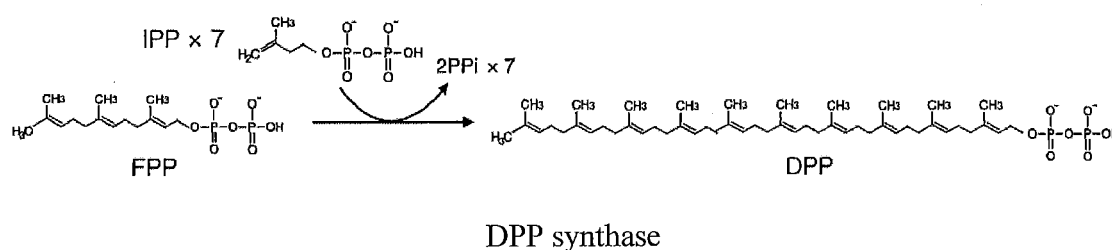
[0097] Mutation A305V identified by a genome analysis based on the constructed conformational model of Enzyme C caused steric hindrance due to repulsion of the atoms with the peripheral amino acid residue, and was deduced to destabilize the conformation of

Enzyme C. A decrease in Enzyme C activity due to the destabilized conformation decreases the amounts of FPP and IPP digested by Enzyme C. As a result, the increase in the amounts of FPP and IPP that can be used for astaxanthin synthesis seemed to increase the amount of astaxanthin produced.

### [0098] 3.1. Enzymatic reaction of decaprenyl diphosphate synthase

Decaprenyl diphosphate synthase is an enzyme that has an activity of catalyzing condensation of FPP with IPP and repeats condensation with IPP to synthesize decaprenyl diphosphate (DPP) from FPP and seven IPPs. The enzymatic reaction of decaprenyl diphosphate synthase is shown below.

#### Chemical formula 2



### [0099] 3.2. Construction of conformational model of Enzyme C

#### (1) Construction of conformational model of Enzyme C by homology modeling

In order to examine the effects of the active site and the mutations in Enzyme C, a conformational model (Figure 16) was constructed by homology modeling using, as a template structure, the conformation of *Rhodobacter capsulatus* (*R. capsulatus*)-derived decaprenyl diphosphate synthase (PDB ID: 3MZV) that had the highest amino acid matching degree (amino acid matching degree of 76.2%) among the conformations registered in PDB, and whose conformation had been determined by X-ray crystallography [4].

The homology modeling was conducted based on the conformation alignment between Enzyme C and the template structure (Figure 17).

#### [0100] (2) Preparation of Enzyme C-substrate complex model

Since the template structure 3MZV is not bonded with a substrate, the conformational data of *Escherichia coli*-derived octaprenyl pyrophosphate synthase (octaprenyl diphosphate synthase) (PDB ID: 3WJN, 3WJO) [11] was used to superimpose

3WJN on the conformational model of Enzyme C and embed FPP as the substrate into the conformational model of Enzyme C, thereby preparing a complex model of Enzyme C and FPP. Next, in the same manner, 3WJO was superimposed and IPP was embedded to prepare a complex model of Enzyme C and FPP and IPP.

Figure 18 shows the template structure and the constructed model structure. Similar to the template structure, Enzyme C forms a homodimer.

### [0101] 3.3. Regarding Enzyme C

#### (1) Conformational comparison

The decaprenyl diphosphate synthase as Enzyme C belongs to the polyprenyl diphosphate synthase family (Pfam PF00348 Polyprenyl synthetase). Figure 19 shows the state of the substrate binding with Enzyme C.

Polyprenyl diphosphate synthase condensates FPP and IPP in the head-to-tail direction (following document [4], the phosphate group side is referred to as the head while the isoprenyl group side is referred to as the tail), thereby synthesizing various polyprenyl diphosphates. By continuing condensation reaction of FPP and IPP, decaprenyl diphosphate synthase extends the prenyl chain deep into the substrate binding site (as represented by an arrow in the right view of Figure 19) to synthesize C50-decaprenyl diphosphate.

[0102] Enzyme C and *R. capsulatus*-derived decaprenyl diphosphate synthase as the template have an amino acid matching degree as high as 76.3% with a conserved active site (Figure 17). RMSD of the conformation by Ca superimposition was 0.063Å, where the two enzymes were very similar. When the matching of the amino acid residues was distinguished by colors, a region having different types of amino acid residues was limited to the molecular surface, the active site and the region binding the substrate entirely consisted of matching amino acid residues (Figure 20).

#### [0103] (2) Comparison by amino acid sequences

Next, the amino acid sequence of Enzyme C was compared with *Paracoccus*-derived decaprenyl diphosphate synthase. *Paracoccus zeaxanthinifaciens* (Q8L1I6)- and *Paracoccus denitrificans* (A1B3M9)-derived amino acid sequences have been submitted to UniProt (<http://www.uniprot.org>) as already known amino acid sequences of *Paracoccus*-derived decaprenyl diphosphate synthases. When they were compared with the amino acid sequence of Enzyme C, the amino acid matching degree was 75.1% (degree

of similarity 89.2%), showing high homology (Table 5).

The amino acid sequence also suggested that Enzyme C was decaprenyl diphosphate synthase. Figure 21 shows the alignments.

Table 5: Comparison of amino acid sequences between Enzyme C and decaprenyl diphosphate synthases

Protein	Organism	UniProt	Identity (similar)
Decaprenyl diphosphate synthase	<i>Paracoccus zeaxanthinifaciens</i>	Q8L1I6	86.2% (95.2%)
	<i>Paracoccus denitrificans</i>	A1B3M9	75.1% (89.2%)

#### [0104] 3.4. Deduction of active site of Enzyme C

In a complex model of Enzyme C and the substrates, the FPP binding site and the IPP binding site as the substrates and the Mg binding site required for catalysis were deduced. The active sites deduced from these results, and the active center and the substrate binding sites of other polyprenyl diphosphate synthases were highly conserved, and thus Enzyme C was deduced to have a reaction mode similar to that of polyprenyl diphosphate synthase.

[0105] The complex model of Enzyme C and the substrates is shown in Figure 22.

Similar to decaprenyl diphosphate synthase as the template structure, Enzyme C was expected to form a homodimer. Figure 22 shows the binding sites of FPP and IPP in chain A (light red). FPP binds to the tunnel region of chain A while FPP and IPP bind to each other in a head-to-tail form with the phosphate group of FPP directing toward the isopentenyl group of IPP. A catalytic reaction occurs between the phosphate group of FPP and the isopentenyl group of IPP in the presence of Mg.

[0106] In order to deduce the active site of Enzyme C, amino acid residues present in the vicinity of FPP and IPP were detected in the Enzyme C-substrate complex model to deduce the interaction between the substrate and the amino acid residues. For FPP, hydrogen bonds between phosphate groups and Arg102, Lys179 and Lys244, and hydrophobic interactions between polyprenyl groups and Ala88, Thr89, His92 and Phe125 were deduced (Figure 23, top image).

For IPP, hydrogen bonds between phosphate groups and Lys54, Arg57, His86 and Arg103, and a hydrophobic interaction between an isopentenyl group and Phe216 were

deduced (Figure 23, center image).

[0107] This catalytic reaction requires Mg. As Mg binding sites, two DDXXD motifs are known in a known polyprenyl diphosphate synthase [11]. Although the coordinates of Mg are not determined in the template structure, Asp93, Asp94, Asp97 and Asp220, and Asp221 and Asp224 corresponding to the DDXXD motifs exist near phosphate groups in Enzyme C similar to the known Mg binding sites, and thus these amino acid residues were deduced to bind Mg (Figure 23, bottom image).

[0108] The deduced active sites are shown in Tables 6-8. In addition, the binding modes between Enzyme C and FPP and IPP as well as the binding modes between Enzyme C and Mg are shown in Figures 24 and 25, respectively.



Table 6: Amino acid residues predicted to have interactions with FPP and IPP detected in Enzyme C

Substrate		Enzyme_C		interaction
FPP	Polyprenyl	Ala88	side-chain	hydrophobic interaction
		Thr89	side-chain	hydrophobic interaction
		His92	side-chain	hydrophobic interaction
		Phe125	side-chain	hydrophobic interaction
	PO4	Arg102	side-chain	hydrogen bond
		Lys179	side-chain	hydrogen bond
		Lys244	side-chain	hydrogen bond
IPP	Isopentenyl	Phe216	side-chain	hydrophobic interaction
	PO4	Lys54	side-chain	hydrogen bond
		Arg57	side-chain	hydrogen bond
		His86	side-chain	hydrogen bond
		Arg103	side-chain	hydrogen bond

[0109]

Table 7: Amino acid residues predicted to have interaction with Mg detected in Enzyme C

Mg	Enzyme_C		Interaction
Mg 1	Asp93	side-chain	hydrogen bond
	Asp94	side-chain	hydrogen bond
	Asp97	side-chain	hydrogen bond
Mg 2	Asp220	side-chain	hydrogen bond
	Asp221	side-chain	hydrogen bond
	Asp224	side-chain	hydrogen bond

[0110]

Table 8: Amino acid residues of deduced active sites in Enzyme C

binding site	residue
FPP	Ala88
	Thr89
	His92
	Arg102
	Phe125
	Lys179
	Lys244
IPP	Lys54
	Arg57
	His86
	Arg103
	Phe216
Mg	Asp93
	Asp94
	Asp97
	Asp220
	Asp221
	Asp224

[0111] Each of the amino acid residues at the deduced bonds matches that of *R. capsulatus*-derived decaprenyl diphosphate synthase as the template structure (Figure 17). In addition, since the FPP binding site, the IPP binding site and the Mg binding site of Enzyme C retain the binding site of *E. coli*-derived octaprenyl diphosphate synthase used for making the substrate complex model, Enzyme C is considered to take a reaction mode similar to octaprenyl diphosphate synthase.

[0112] In particular, Arg102, Lys179 and Lys244 that bind to the phosphate group of FPP, and Asp93, Asp94, Asp97, Asp220, Asp221 and Asp224 that bind to Mg seem to be important residues that are directly responsible for the activity of catalyzing transfer of the phosphate groups. Experiments of preparing rat and yeast FPP synthase mutants reported

that amino acid residues corresponding to Arg102 and Arg103 as phosphate binding sites and Asp94, Asp97, Asp220, Asp221 and Asp224 as Mg binding sites of Enzyme C were important for enzymatic activity, and that these amino acid residues were highly conserved among polyprenyl diphosphate synthases [14, 15].

[0113] In addition, in FPP synthase, side chains of Phe and Gln were found to be important for substrate binding through preparation of mutants of the amino acid residues corresponding to Phe216 and Gln217 of Enzyme C [16]. Phe216 and Gln217 also existed in the region of the active site in the complex model of Enzyme C and thus were predicted to be important for the activity.

#### [0114] 3.5. Effect of mutation A305V in Enzyme C

In order to deduce the effect of mutation from Ala305 to Val identified in Enzyme C on the conformation, a conformational model of mutant A305V enzyme of Enzyme C was prepared.

##### [0115] (1) Preparation of mutant model by A305V single residue substitution

In order to observe the effect of substitution from Ala to Val, a model was first prepared by fixing the conformation of the rest of the amino acid residues (assuming it to be a rigid body) and allowing a single residue substitution from Ala305 to Val, to compare the model with the wild type.

[0116] Figure 26 shows conformational models of the wild-type and mutant A305V enzymes. Ala305 in the wild type was present in the  $\alpha$ -helix (pink), where the side chain provided packing with the amino acid residue of the adjacent  $\alpha$ -helix (cyan) via a hydrophobic interaction.

[0117] Figure 27 shows the structure of the amino acid residue adjacent to Ala305.

The carbon atom of the methyl group as the side chain of Ala305 makes contact with the peripheral amino acid residues Tyr208, Ala211, His301 and Ala302, where the interatomic distances between the carbons were all less than 4.0Å. The substitution from Ala305 to Val results in an addition of two methyl groups to the side chain. In the conformational model including the substitution from Ala305 to Val, the interatomic distances between the carbon of the methyl group of the side chain of Val and the carbons of Tyr208 and Ala211 were 2.42Å and 2.47Å, respectively.

[0118] The lower limit of the contact distance between carbons via a noncovalent bond is

2.9Å. Since the interatomic distances of the carbons measured with respect to Val305 was lower than this value, these carbon atoms would repel and cause steric hindrance with the peripheral amino acid residues. In a space-filling representation, the interatomic distances were shorter than the van der Waals radii, confirming repulsion between the atoms. Accordingly, A305V mutation was predicted to induce steric hindrance in Enzyme C due to repulsion of the atoms, which leads to destabilization of the conformation.

[0119] (2) Change in intramolecular energy due to mutation A305V

In order to examine the destabilization of the structure caused by A305V, the intramolecular energies of the wild-type and mutant A305V Enzyme C were computed in terms of unit kilojoules/mol (KJ/mol) from a set of the interatomic bond length, the bond angle, the torsion, the bond energy and the like.

[0120] As a result, mutant A305V enzyme gave an intramolecular energy of -16,295 (KJ/mol) while the wild type gave -17,912 (KJ/mol), showing an increase of 9.02% and confirming destabilization of the structure. This increase in the intramolecular energy were particularly observed at the amino acid residues of Tyr208, Ala211 and Val305, and thus repulsion caused by these amino acid residues seemed to result the increase in the intramolecular energy. Comparison of the intramolecular energies between the wild-type and mutant A305V Enzyme C is shown in Figure 28.

[0121] (3) Structural change caused by mutation A305V

Next, an energy minimization calculation was performed on the mutant model so as to examine if repulsion due to A305V can be avoided by moving the conformation of the peripheral amino acid residues. As a result, it was found that the conformation of the amino acid residues of the  $\alpha$ -helix adjacent to the  $\alpha$ -helix in which Val305 exists needs to be moved in order to avoid repulsion of the side chain of Val305 while accepting the side chain. Specifically, it was predicted that this structural change not only affects the side chain but also the main chain, and thus the original packing between the  $\alpha$ -helices does not occur in the mutant A305V enzyme and the structure surrounding the two helices was destabilized. Furthermore, the active sites that bond substrate IPP and Mg (Phe216, Gln217, Asp220) are present in the  $\alpha$ -helix adjacent to mutation A305V. The structural change in the main-chain structure of the  $\alpha$ -helix as the basis of the active sites dislocates the amino acid residues of the active sites, and this would have an effect on the substrate bonds and activity themselves

(Figure 29, right).

[0122] Figure 29 shows the structures around Ala305 and Val305 in the wild type and the mutant A305V following energy minimization calculation, while Figures 30, 31 and 32 show the structural changes in mutant A305V.

[0123] (4) Decrease in Enzyme C activity due to structural destabilization

Thus, from the conformational model of A305V, occurrence of the steric hindrance with the peripheral amino acid residues, the increase in the intramolecular energy and the structural change in the adjacent two  $\alpha$ -helices were deduced, suggesting that mutation A305V destabilizes the conformation of Enzyme C. Conformational destabilization caused by point mutation has also been reported frequently in genetic disorders and the like. For example, a protein that has a destabilization-inducing mutation may be deactivated in shorter time than usual for being unable to maintain the original conformation in a solvent, and the mutant protein that cannot give original packing in the cell is eliminated by the function of the cell itself. Accordingly, the mutant A305V enzyme that does not have the original activity and that has an unstable structure may be eliminated in the bacterium, which is deduced to consequently decrease the Enzyme C activity in the bacterium.

[0124] 3.6. Effect of mutation A305V in Enzyme C on astaxanthin synthesis pathway

Decaprenyl diphosphate synthase is one of the enzymes in the coenzyme C10 (CoQ10) synthesis pathway. *Paracoccus zeaxanthinifaciens*- or *Paracoccus denitrificans*-derived decaprenyl diphosphate synthase that was confirmed to be highly homologous with Enzyme C this time is found to be an enzyme required for CoQ10 production [Japanese Patent Application Publication No. 2005-211020, Japanese Patent Application Publication No. 2006-517794]. FPP and IPP as the substrates of decaprenyl diphosphate synthase also serve as substrates of geranyl-geranyl diphosphate (GGPP) synthase CrtE in the astaxanthin synthesis pathway. Therefore, in a usual *Paracoccus* cell, decaprenyl diphosphate synthase and CrtE are considered to compete for substrates FPP and IPP.

[0125] Mutation A305V identified in Enzyme C was deduced to decrease the Enzyme C activity by destabilizing the molecular conformation. The decreased Enzyme C activity means decreased amounts of the substrates FPP and IPP used, which would increase the

amounts of FPP and IPP that can be used in the astaxanthin synthesis pathway.

[0126] CrtE synthesizes one molecule of GGPP from one molecule of FPP and one molecule of IPP. Meanwhile, Enzyme C requires one molecule of FPP and seven molecules of IPP to synthesize one molecule of decaprenyl diphosphate. In terms of IPP, Enzyme C digests IPP seven times as much as CrtE in a single reaction. Thus, the decrease in the Enzyme C activity seems very effective for increasing the IPP supplied to the astaxanthin synthesis pathway.

[0127] As a result, the amount of the synthesized astaxanthin is deduced to increase significantly (Figure 33).

[0128] REFERENCES

1. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T. *Nucleic Acids Res.* 2014 ;42:W252-8.

2. Automated comparative protein structure modeling with SWISS-MODEL and Swiss- PdbViewer: A historical perspective. Guex, N., Peitsch, M.C., Schwede, T. *Electrophoresis*, (2009). 30(S1), S162-S173.

3. Crystal structure of 1-deoxy-D-xylulose 5-phosphate synthase, a crucial enzyme for isoprenoids biosynthesis. Xiang S, Usunow G, Lange G, Busch M, Tong L. *J Biol Chem.* 2007 26;282(4): 2676-82.

4. Prediction of function for the polyprenyl transferase subgroup in the isoprenoid synthase superfamily. Wallrapp FH, Pan JJ, Ramamoorthy G, Almonacid DE, Hillerich BS, Seidel R, Patskovsky Y, Babbitt PC, Almo SC, Jacobson MP, Poulter CD. *Proc Natl Acad Sci USA.* 2013 26;110(13):E1196-202.14;107(50):21337-42.

6. Snapshot of a key intermediate in enzymatic thiamin catalysis: crystal structure of the alpha-carbanion of (alpha,beta-dihydroxyethyl)-thiamin diphosphate in the active site of transketolase from *Saccharomyces cerevisiae*. Fiedler E, Thorell S, Sandalova T, Golbik R, König S, Schneider G. *Proc Natl Acad Sci USA.* 2002 22;99(2):591-5.

7. Examination of substrate binding in thiamin diphosphate-dependent transketolase by protein crystallography and site-directed mutagenesis. Nilsson UI,

Meshalkina L, Lindqvist Y, Schneider G. J Biol Chem. 1997 17;272(3):1864-9.

8. Catalytically Important Residues in E. coli 1-Deoxy-D-Xylulose 5-Phosphate Synthase. Jordi Querol-Audí, Albert Boronat, Josep J. Centelles, Santiago Imperial Journal of Biosciences and Medicines, 2014, 2, 30-35

9. Functional effect of grapevine 1-deoxy-D-xylulose 5-phosphate synthase substitution K284N on Muscat flavour formation. Battilana J1, Emanuelli F, Gambino G, Gribaudo I, Gasperi F, Boss PK, Grando MS. J Exp Bot. 2011;62(15):5497-508.

10. Feedback Inhibition of Deoxy-D-xylulose-5-phosphate Synthase Regulates the Methylerythritol 4-Phosphate Pathway. Banerjee A1, Wu Y, Banerjee R, Li Y, Yan H, Sharkey TD. J Biol Chem. 2013 7;288(23):16926-36.

11. Crystal structures of ligand-bound octaprenyl pyrophosphate synthase from Escherichia coli reveal the catalytic and chain-length determining mechanisms. Han X, Chen CC, Kuo CJ, Huang CH, Zheng Y, Ko TP, Zhu Z, Feng X, Wang K, Oldfield E, Wang AH, Liang PH, Guo RT, Ma Y. Proteins. 2015;83(1):37-45.

14. Effect of site-directed mutagenesis of conserved aspartate and arginine residues upon farnesyl diphosphate synthase activity. Joly A, Edwards PA. J Biol Chem. 1993 25;268(36):26983-9.

15. Yeast farnesyl-diphosphate synthase: site-directed mutagenesis of residues in highly conserved prenyltransferase domains I and II. Song L1, Poulter CD. Proc Natl Acad Sci USA. 1994 12;91(8): 3044-8.

16. Significance of Phe-220 and Gln-221 in the catalytic mechanism of farnesyl diphosphate synthase of Bacillus stearothermophilus. Koyama T1, Tajima M, Nishino T, Ogura K. Biochem Biophys Res Commun. 1995 7;212(2):681-6.

## CLAIMS

1. A mutant carotenoidogenic bacterium belonging to the genus *Paracoccus*, comprising a combination of:

- (i) any one of genes selected from the group consisting of (a)-(e) below; and
- (ii) any one of genes selected from the group consisting of (f)-(j) below :

(a) a gene encoding a protein comprising a mutant amino acid sequence in which at least glycine of the 225th amino acid residue in the amino acid sequence of SEQ ID NO:2 of 1-deoxy-D-xylulose 5-phosphate synthase (DXS) has been substituted with aspartic acid;

(b) a gene encoding a protein comprising a mutant amino acid sequence in which glycine of the 225th amino acid residue in the amino acid sequence of SEQ ID NO:2 of DXS is substituted with aspartic acid and in which one to several amino acid residues other than the 225th amino acid residue have been deleted, substituted or added, and that has DXS activity;

(c) a gene consisting of DNA comprising the nucleotide sequence of SEQ ID NO:5;

(d) a gene comprising DNA that hybridizes with DNA consisting of a nucleotide sequence complementary to the gene of (b) above under stringent conditions, and that encodes a protein having increased DXS activity, wherein the nucleotide sequence encoding the 225th amino acid residue in the gene always hybridizes as a codon for an aspartic acid residue;

(e) a gene comprising DNA that hybridizes with DNA consisting of a nucleotide sequence complementary to DNA comprising the nucleotide sequence of SEQ ID NO:5 under stringent conditions, and that encodes a protein having increased DXS activity, wherein the nucleotide sequence of positions 673-675 in the gene always hybridizes as codon for an aspartic acid residue;

(f) a gene encoding a protein comprising a mutant amino acid sequence in which at least alanine of the 305th amino acid residue in the amino acid sequence of SEQ ID NO: 4 of decaprenyl diphosphate synthase (DPS) has been substituted with valine;

(g) a gene encoding a protein comprising a mutant amino acid sequence in which



alanine of the 305th amino acid residue in the amino acid sequence of SEQ ID NO:4 of DPS is substituted with valine and in which one to several amino acid residues other than the 305th amino acid residue have been deleted, substituted or added, and that has reduced DPS activity;

(h) a gene consisting of DNA comprising the nucleotide sequence of SEQ ID NO:7;

(i) a gene comprising DNA that hybridizes with DNA consisting of a nucleotide sequence complementary to the gene of (g) above under stringent conditions, and that encodes a protein having reduced DPS activity, wherein the nucleotide sequence encoding the 305th amino acid residue in the gene always hybridizes as a codon for a valine residue; and

(j) a gene comprising DNA that hybridizes with DNA consisting of a nucleotide sequence complementary to DNA comprising the nucleotide sequence of SEQ ID NO:7 under stringent conditions, and that encodes a protein having reduced DXS activity, wherein the nucleotide sequence of positions 913-915 in the gene always hybridizes as codon for a valine residue.

2. The bacterium according to Claim 1, which has acquired carotenogenic capacity that is higher than the carotenogenic capacity of a carotenoidogenic bacterium without the genes encoding the proteins comprising the mutant amino acid sequences.

3. The bacterium according to Claim 2, which has acquired carotenogenic capacity that is at least 5 times or more the carotenoid production amount of a carotenoidogenic bacterium without the genes encoding the proteins comprising the mutant amino acid sequences.

4. The bacterium according to Claim 1, wherein the bacterium belonging to the genus *Paracoccus* is strain E-396.

5. The bacterium according to any one of Claims 1-4, wherein the carotenoid is astaxanthin.

6. A method for producing a carotenoid, comprising culturing the bacterium according to any one of Claims 1-5, and collecting the carotenoid from the resulting cultured product.
7. The method according to Claim 6, wherein the carotenoid production amount is at least 5 times or more the carotenoid production amount of a carotenoidogenic bacterium without the genes encoding the proteins comprising the mutant amino acid sequences.
8. The method according to Claim 6 or 7, wherein the carotenoid is astaxanthin.
9. A recombinant vector comprising the combination of genes as described in claim 1.
10. A transformant comprising the recombinant vector according to Claim 9.
11. A method for producing a carotenoid, comprising culturing the transformant according to Claim 10, and collecting a carotenoid from the resulting cultured product.

Fig. 1

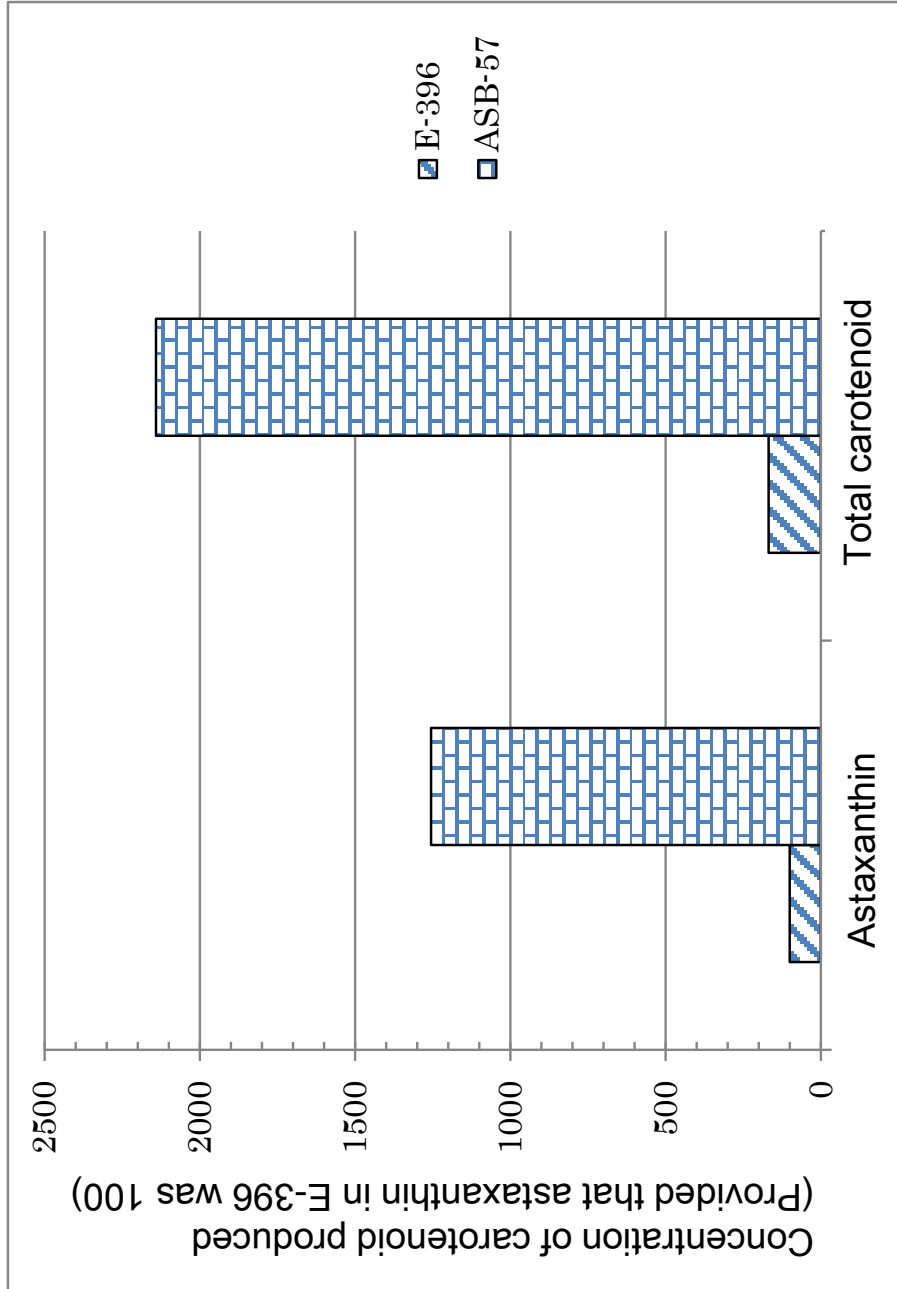


Fig. 2

Protein	1-deoxy-D-xylulose-5-phosphate synthase		
UniProt	Q9RUB5 (DXS_DEIRA)	EC:2.2.1.7	
Organism	<i>Deinococcus radiodurans</i>		
Pfam	PF13292	DXP_synthase_N	
PDB ID	2O1X	Resolution: 2.90[Å]	dimer
Identity	44.1%	7-630	

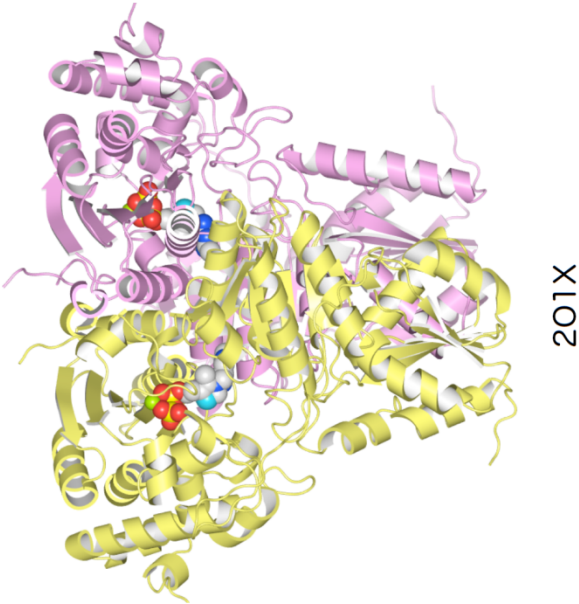


Fig. 3A

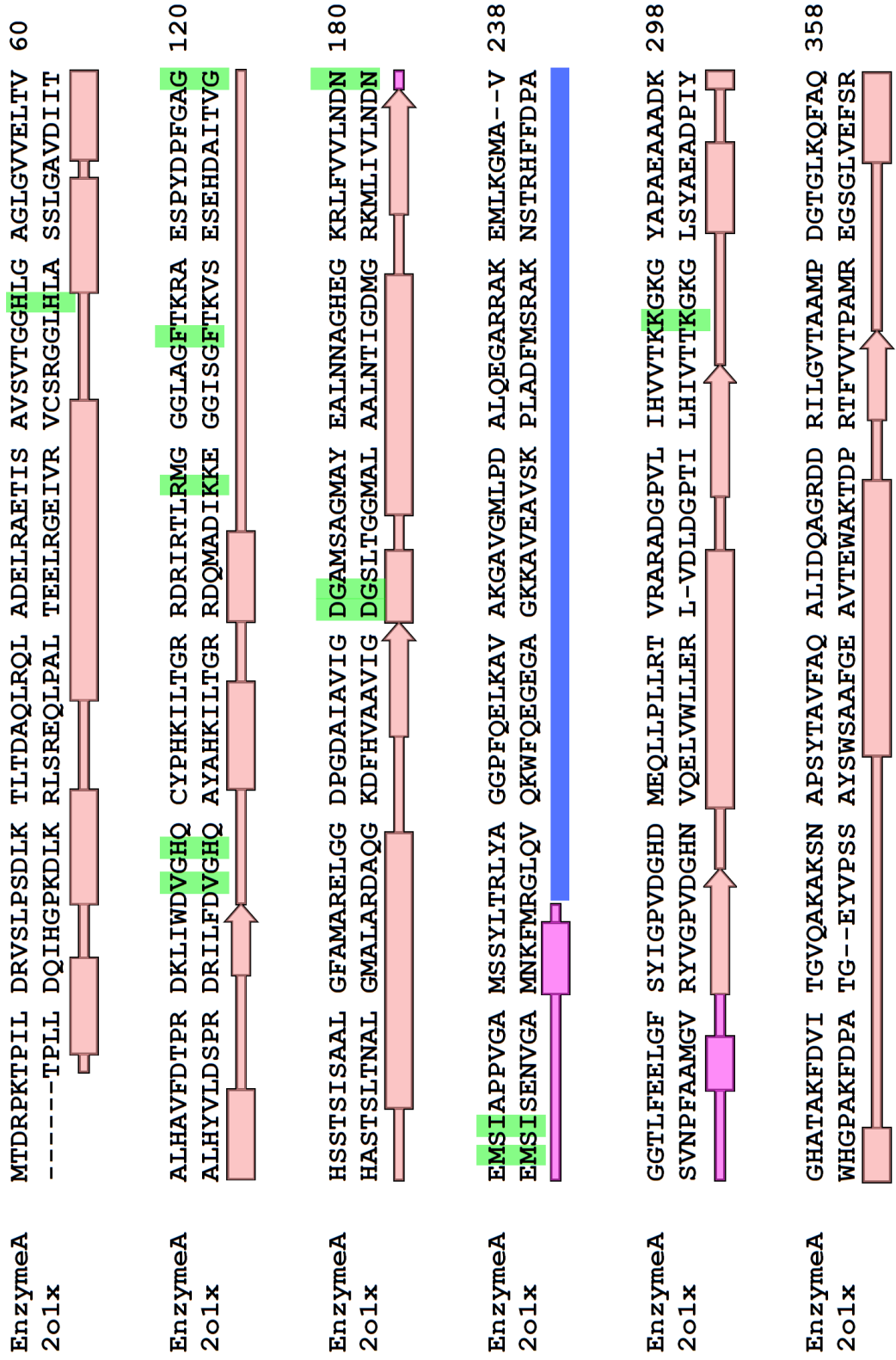


Fig. 3B



Fig. 4

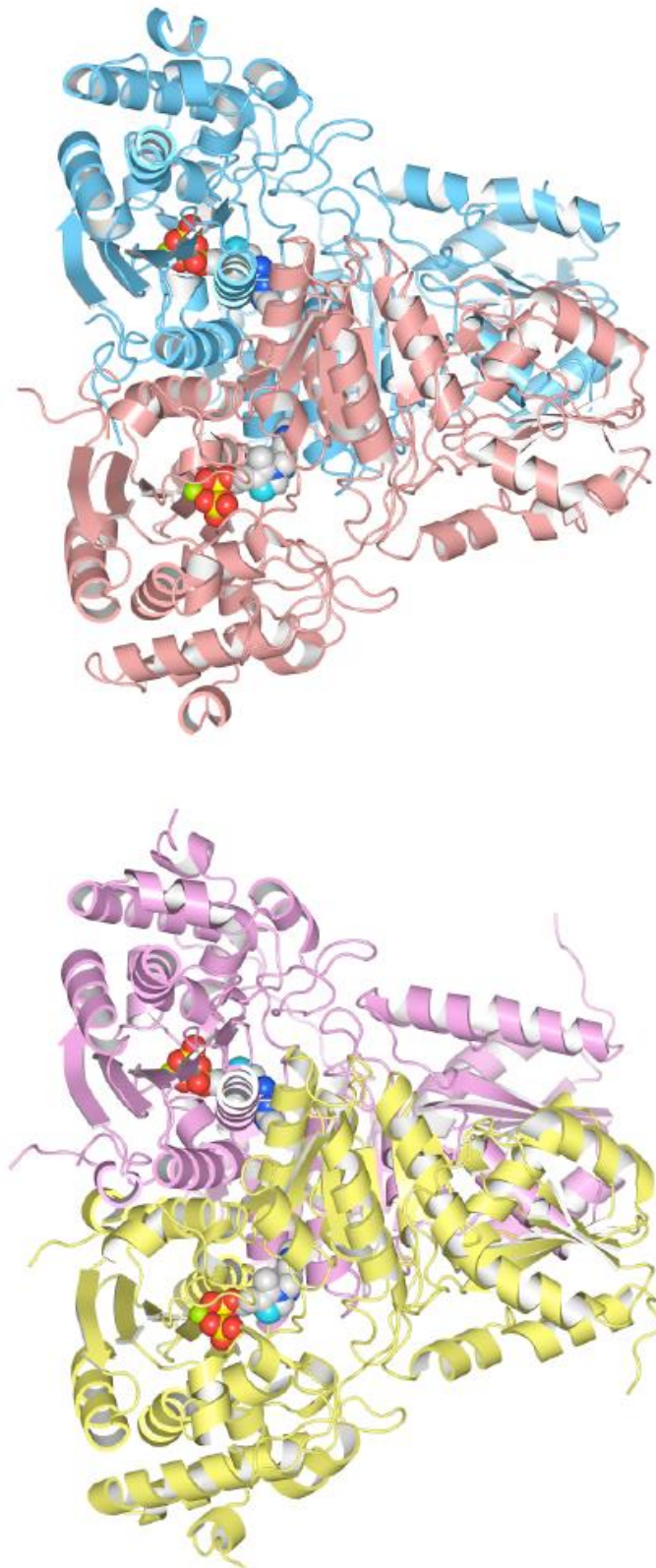


Fig. 5

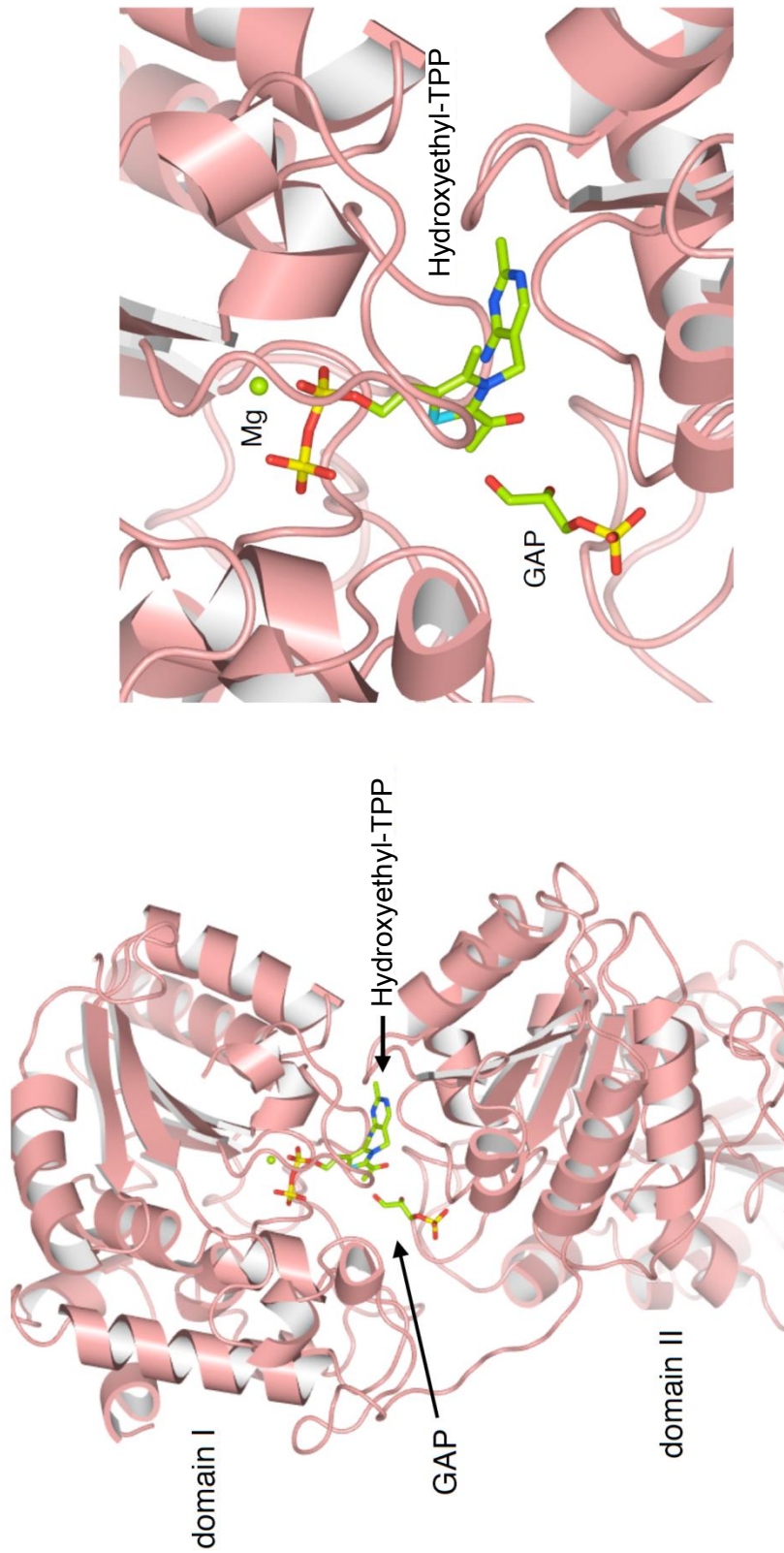




Fig. 6

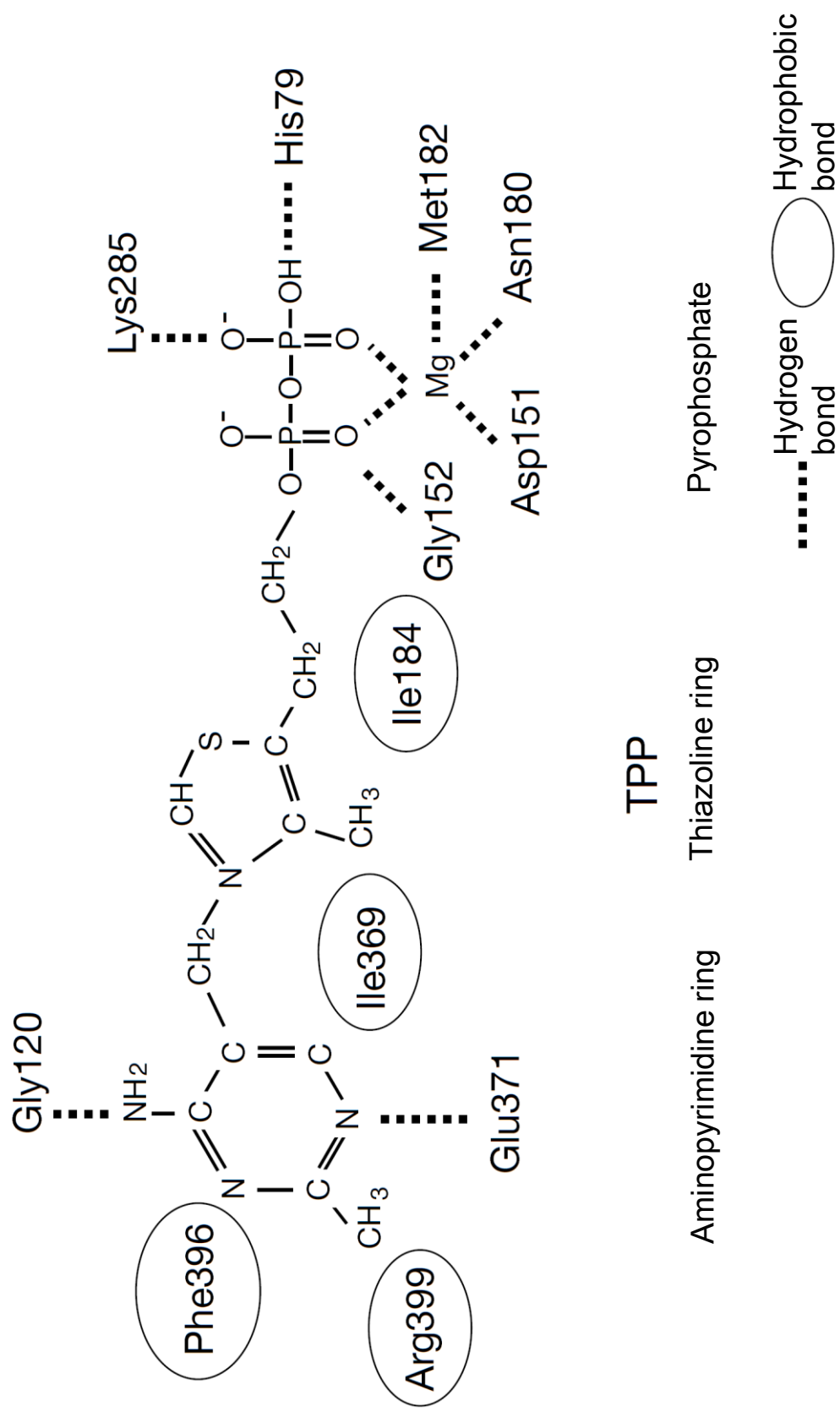


Fig. 7

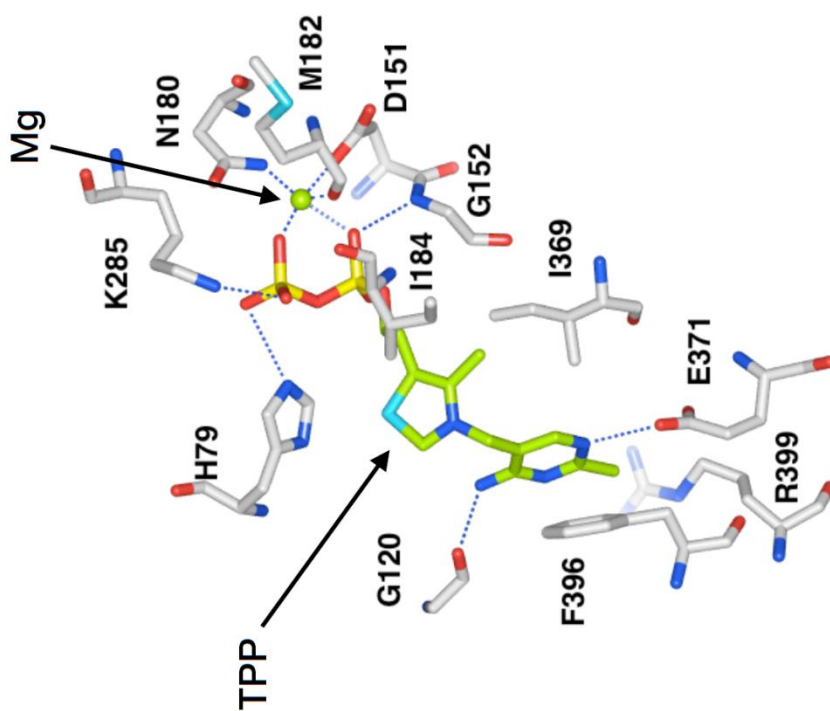
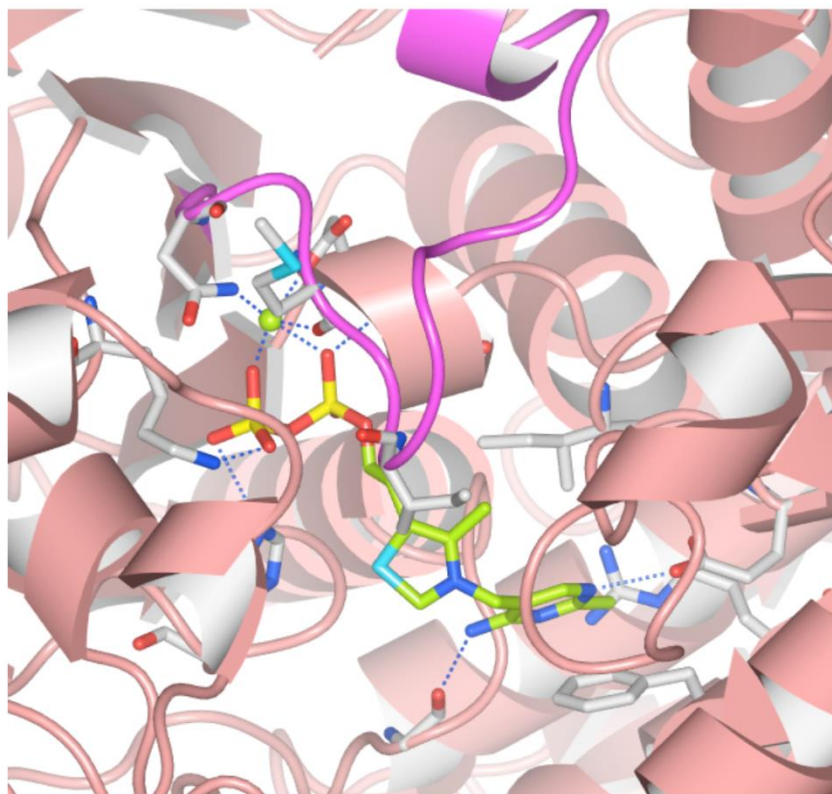


Fig. 8

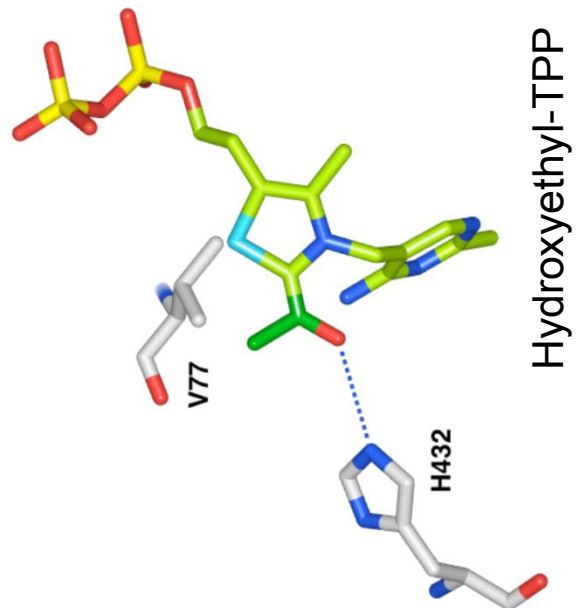
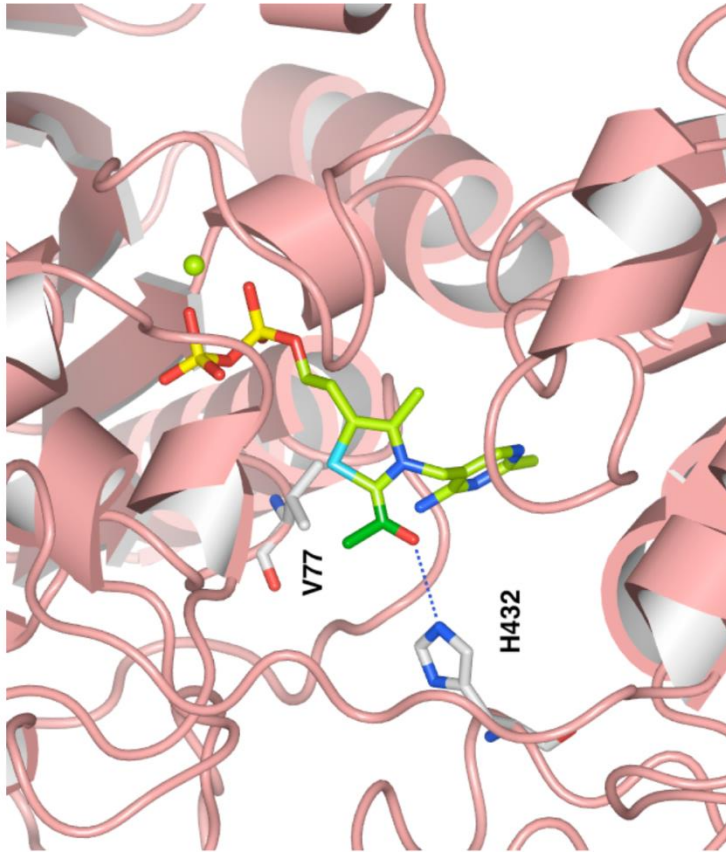


Fig. 9

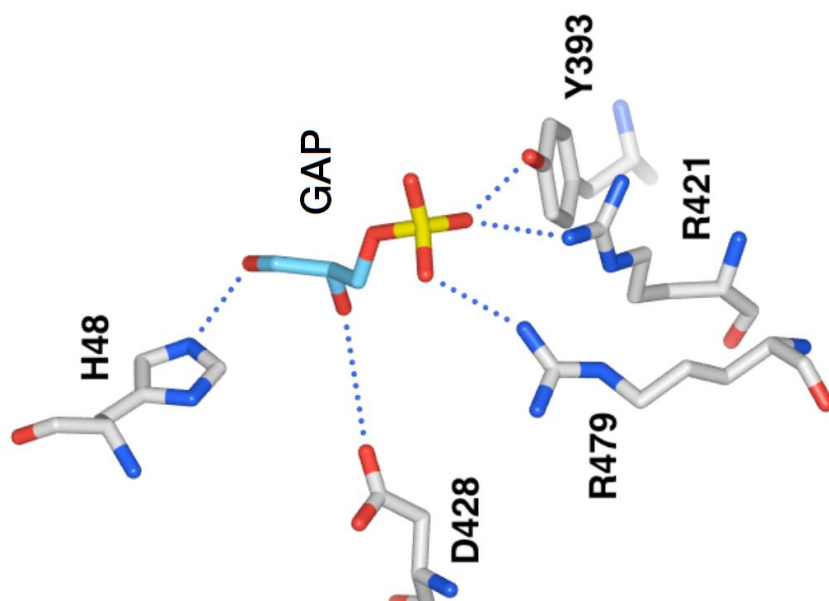
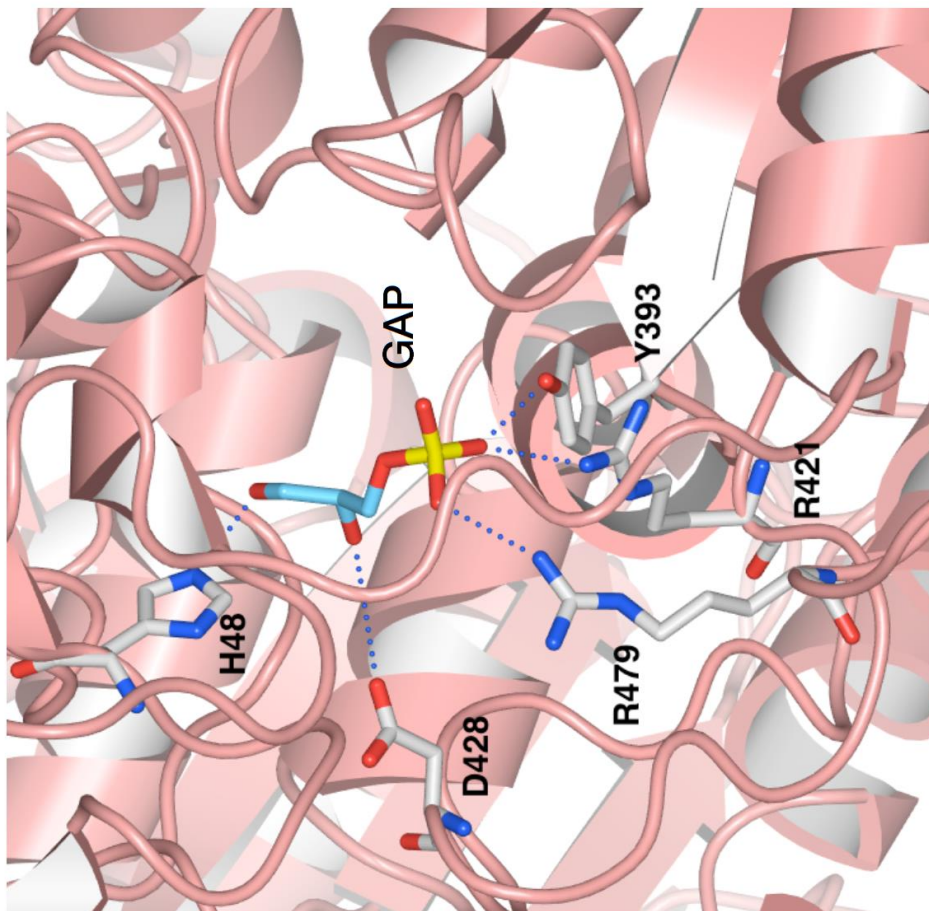


Fig. 10

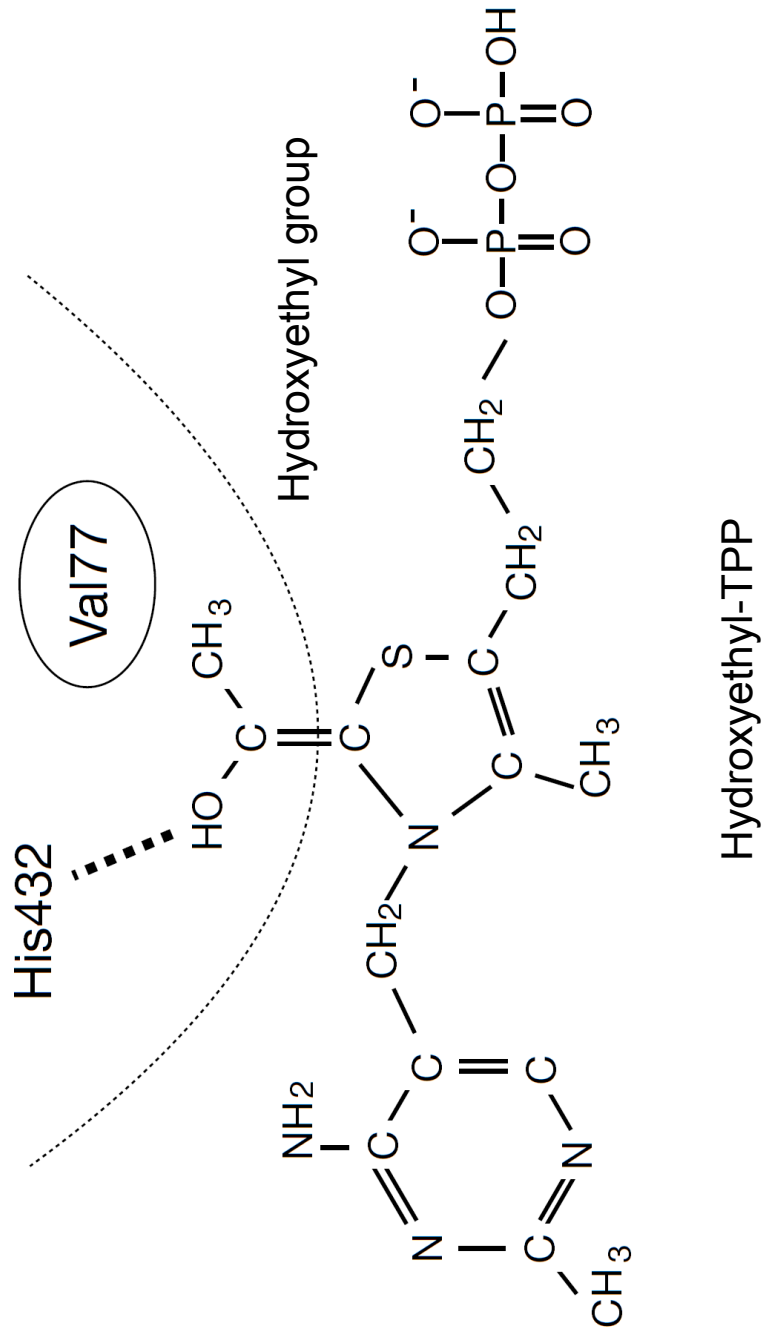
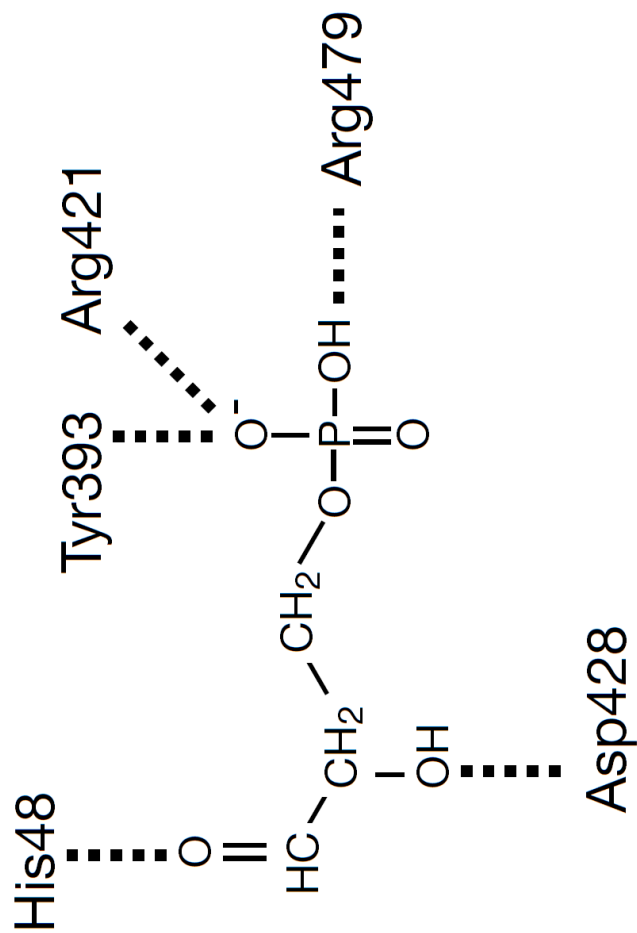


Fig. 11



Glyceraldehyde 3-phosphate

..... Hydrogen bond

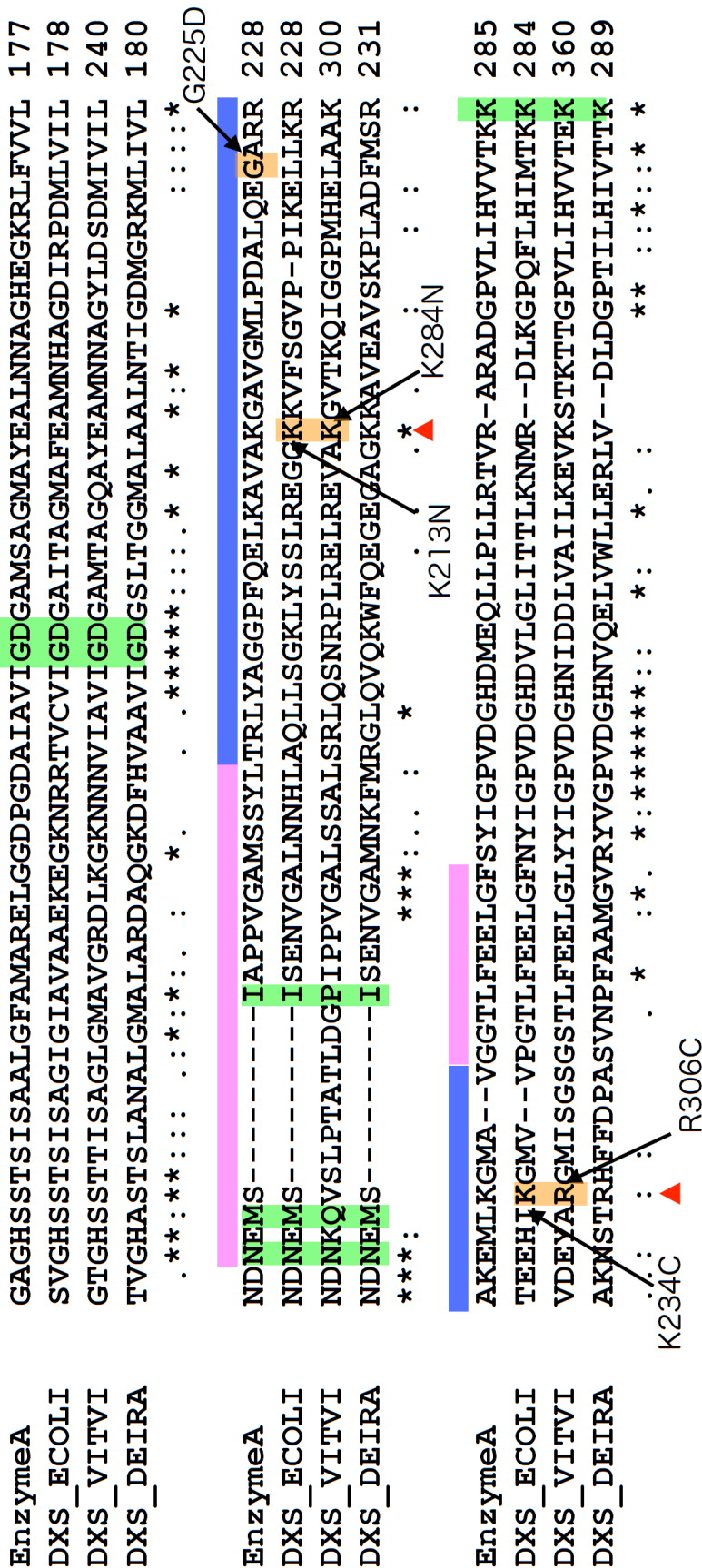




Fig. 13

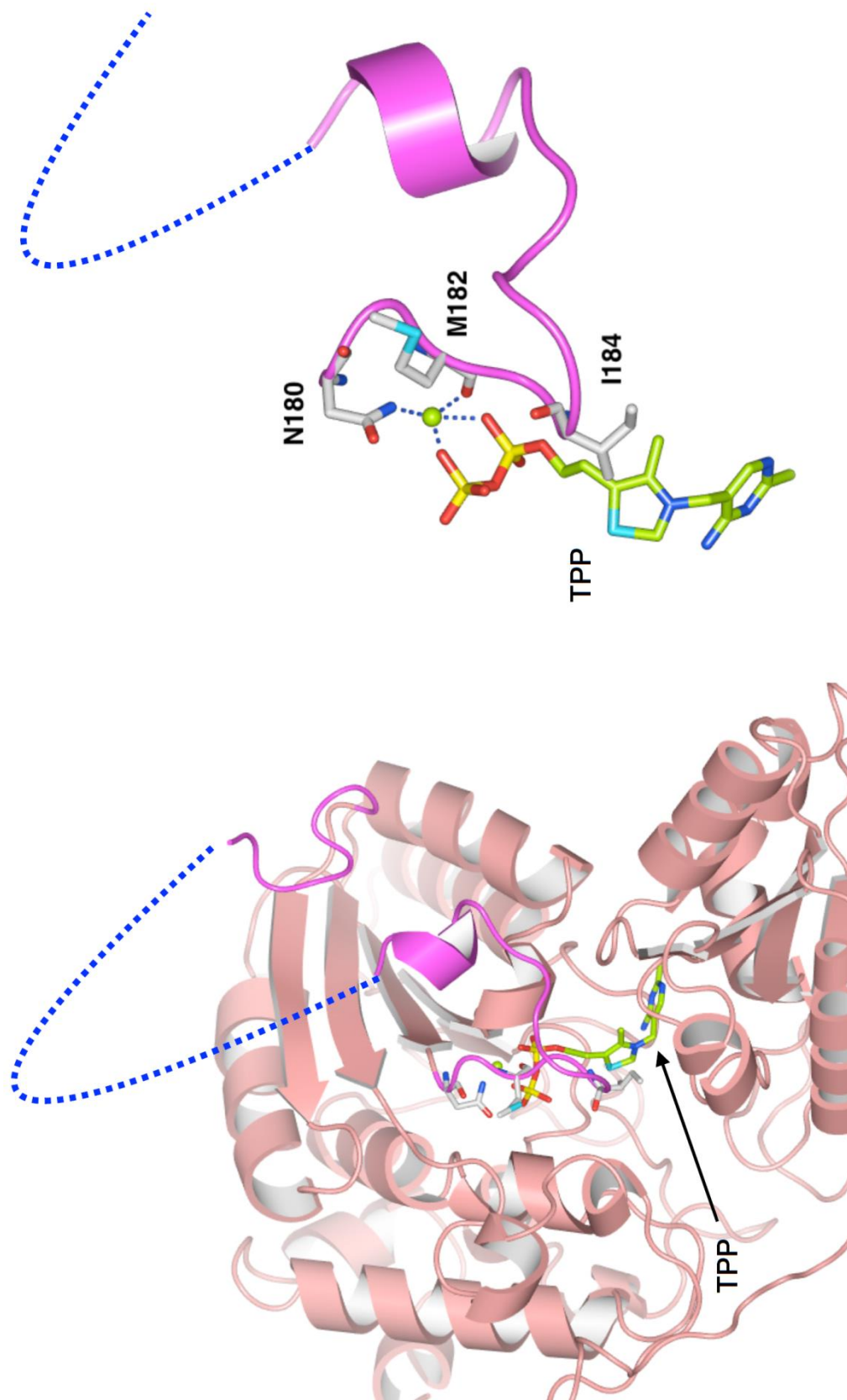




Fig. 14

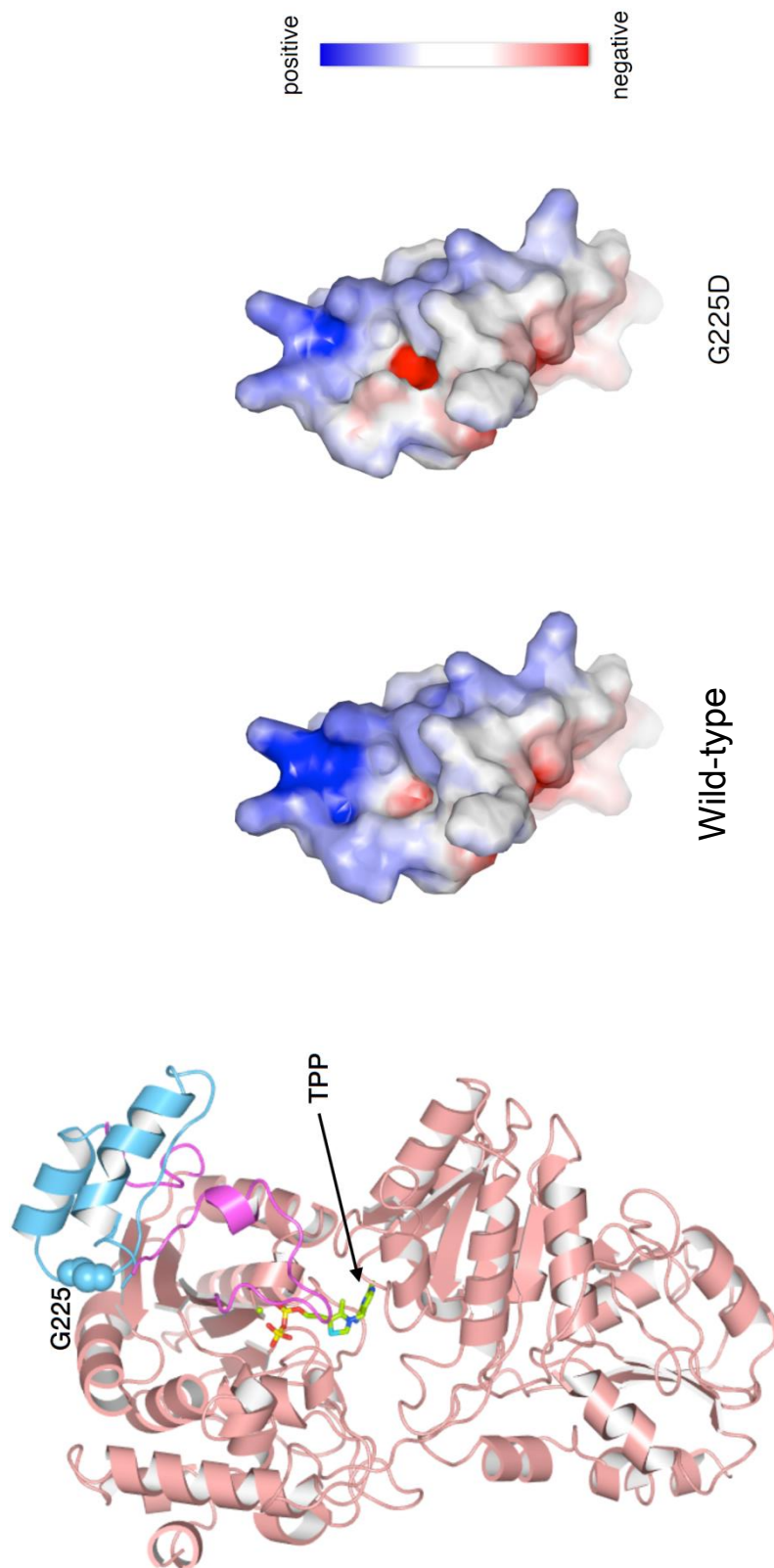


Fig. 15

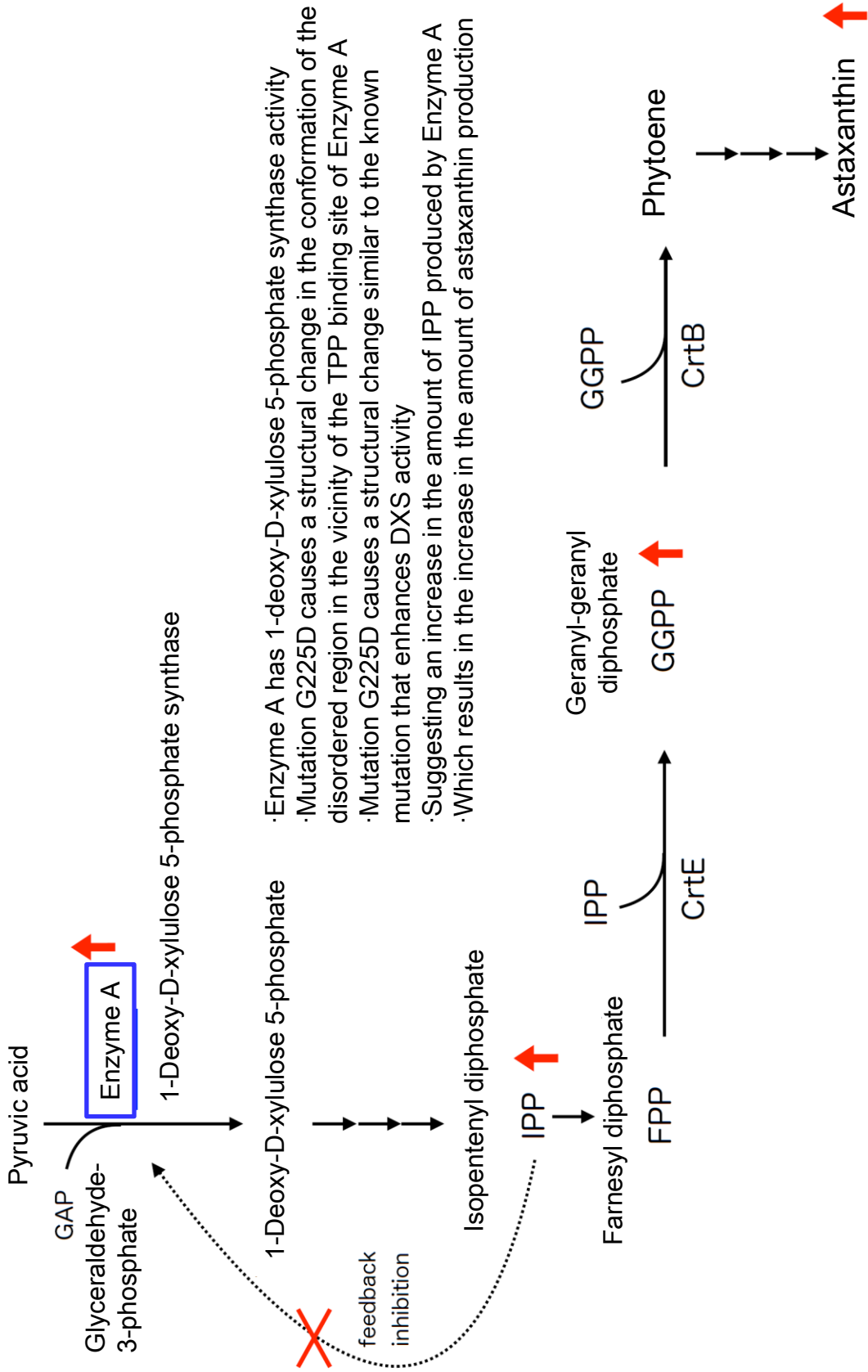
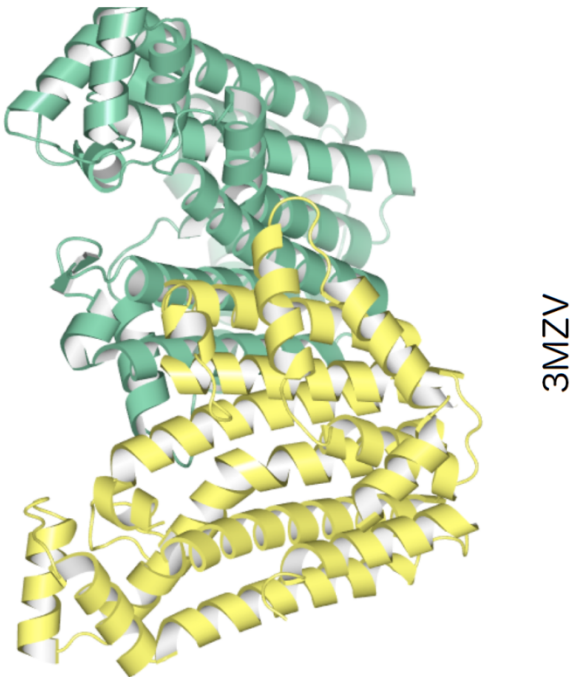


Fig. 16

Protein	Decaprenyl diphosphate synthase		
UniProt	Q07D00 (Q07D00_RHOCA)	EC:2.5.1.91	
Organism	<i>Rhodobacter capsulatus</i>		
Pfam	PF00348	Polyprenyl synthetase	
PDB ID	3MZV	Resolution: 1.90[Å] dimer	
Identity	76.2%	10-332	



EnzymeC 3mzv	MTVQDNVRKP MDRLSEALTA EMEAVNALIR DRMASRHAPR IPEVTAHLIE AGGKRLRPML 60 MSLDDKATKP HDRLAQALAE DMAAVNALIR ERMSSEHAPR IPEVTAHLIE AGGKRLRPML
EnzymeC 3mzv	TLAAAKLLGY PGPWHVHLAA TVEFIHTATL LHDDVVDESA QRRGRPTANL LWDNKSSVLV 120 TLAAARLVGY GGFHVHLAA TVEFIHTATL LHDDVVDESR QRRGRPTANL LWDNKSSVLV
EnzymeC 3mzv	GDYLFARSEQ LMVEPGNLR TLEILANASAT IAEGEVLQLT AAQDLATDES VYLQVVRGKT 180 GDYLFARSEQ LMTDTGNMRV MEILANASAV IAEGEVLQLT AAQNLATTED IYLRVIRGKT
EnzymeC 3mzv	AALFSAATEV GGVIAGAPDD QVQALFDYGD ALGISFQIVD DLLDYGGATE TIGKNVGDDF 240 AALFSAATEV GGIIGGAPED QVQALFDYGD ALGIAFQIVD DLLDYGGKSA EIGKNTGDDF
EnzymeC 3mzv	RERKLTLPVI KAIKADAEAE RAFWTRTIEA GDQRDGDLEH ALSLLARHGA MEAARADALA 300 RERKLTMPVI KAVALADEAE RAFWKRVIEK GDQQDGDLEH AMALMTKHGT LEATRLAAIG
EnzymeC 3mzv	HAARVRAALQ VIPAHPIRM LADLADFVVS RV 332 WTD TARKALA KLPDHPLRQM LDDLADYVVE RV

Fig. 18

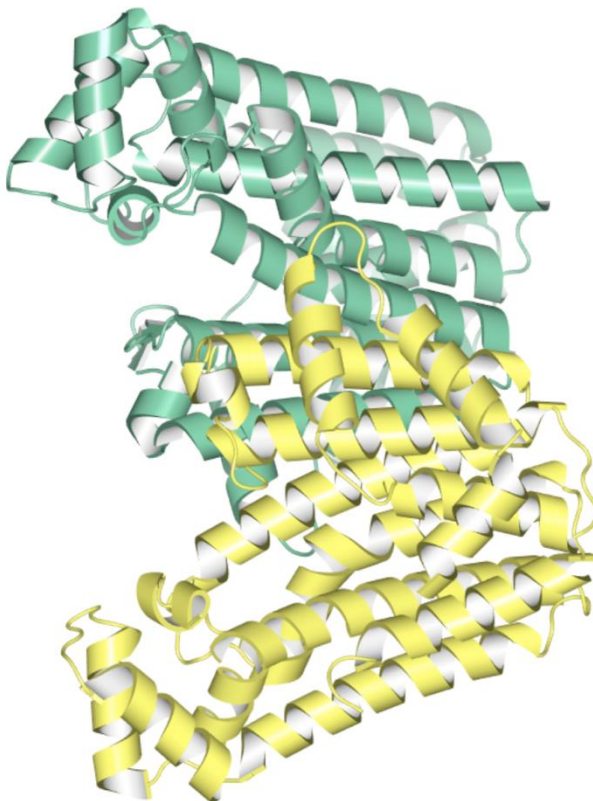
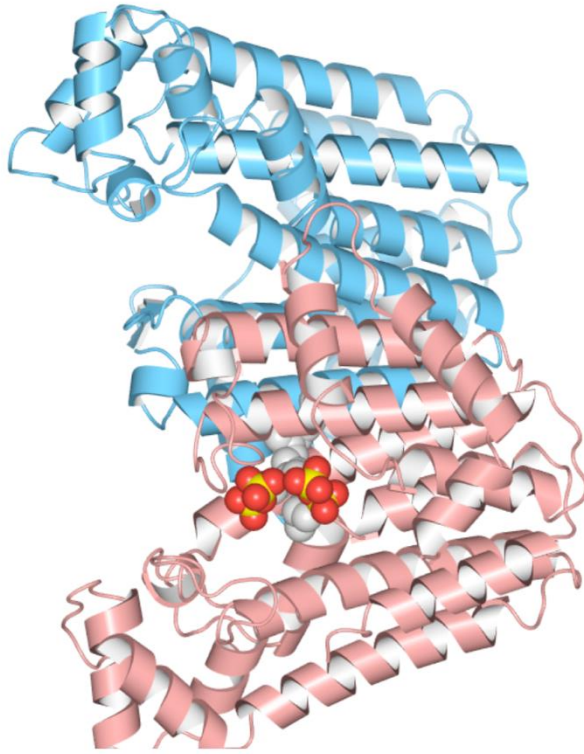


Fig. 19

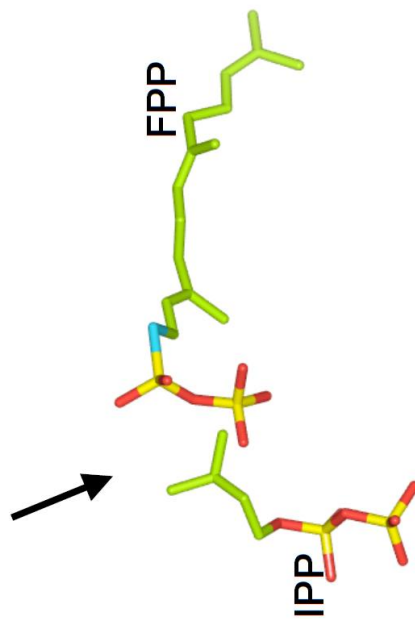
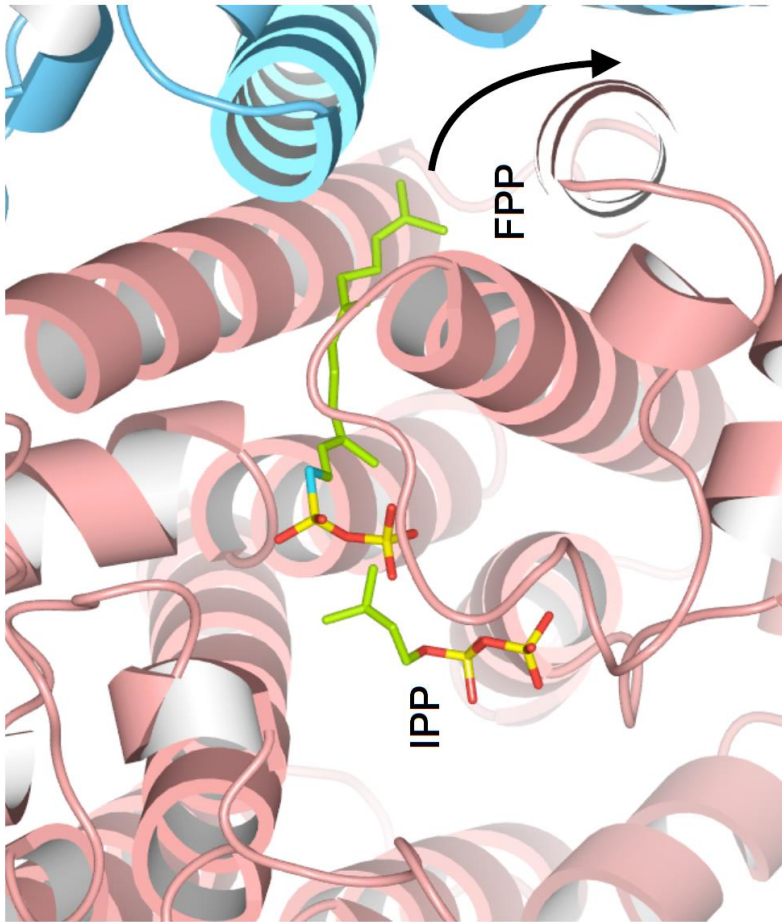
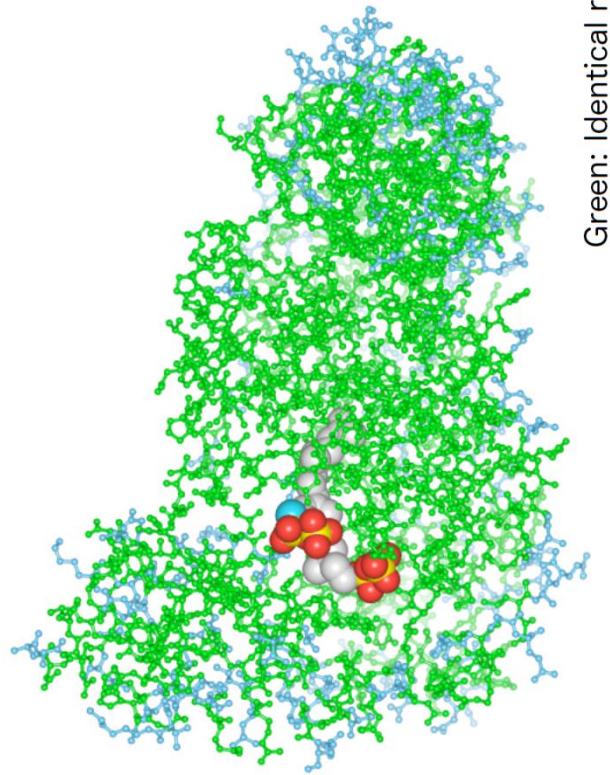
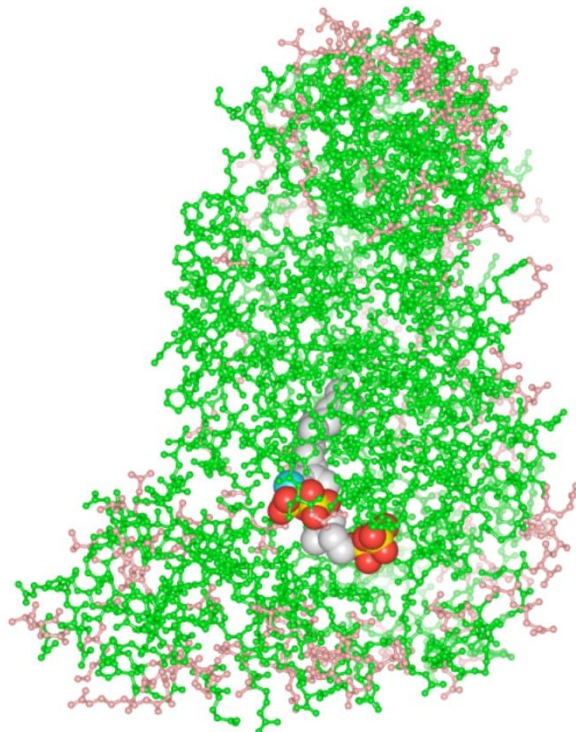




Fig. 20



Green: Identical residues



```

>>tr|Q8L1I6|Q8L1I6_PARZE                               (333 aa)
initn: 1887 initl: 1859 opt: 1866 Z-score: 2009.0 bits: 380.1 E(5): 2.1e-109
Smith-Waterman score: 1866; 86.2% identity (95.2% similar) in 333 aa overlap (1-333:1-333)

      10      20      30      40      50      60
EnzymC  MTVDNVRKPMDRLSEALTAEMEAVNALIRDRMASRHAPRIPEVTAHLIEAGGKRLRPML
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Q8L1I6  MNVQEDVRKPLDRLAEALAPEMEAVNALIRERMASRHAPRIPEVTAHLIEAGGKRLRPML

      70      80      90     100     110     120
EnzymC  TLAAAKLLGYPGPWHVHLAATVEFIHTATLLHDDVVDESAQRGRPTANLLWDNKSSVIV
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Q8L1I6  TLAAAKLLGYGGPYHVHLAATVEFIHTATLLHDDVVDESQRGRPTANLLWDNKSSVIV

      130     140     150     160     170     180
EnzymC  GDYLFARSFQLMVEPGNLRLEILANASATIAEGEVLQLTAAQDLATDESIVLQVVRGKT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Q8L1I6  GDYLFARSFQLMVEPGSMRTEILSNAATIAEGEVLQLTAAQDLATNEDIYVLRGKT

      190     200     210     220     230     240
EnzymC  AALFSAATEVGGVIAGAPDDQVQALFDYGDALGISFQIVDDLDDYGGATETIGKNVGDDF
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Q8L1I6  AALFSAATEVGGVIAGVPDAQVRALFDYGDALGIAFQIVDDLDDYGGTAEAIKNTGDDF

      250     260     270     280     290     300
EnzymC  RERKLTLPVIKAIKADAEERAFWTRTIEAGDQDGDLEHALSLARHGAMEAARADALA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Q8L1I6  RERKLTLPVIKAVARATPEERAFWSRTIEKGDQKDGDLLEHALELLARHGAMADARRDALD

      310     320     330
EnzymC  HAARARALQVLPAPHPIRDMLADLADFVVSVA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Q8L1I6  WAARARASLIQLPEHPIRDMLSLADLADFVVERIA

```

EnzymC : Enzyme\_C

Q8L1I6 : Decaprenyl diphosphate synthase (Paracoccus zeaxanthini faciens)



Fig. 22

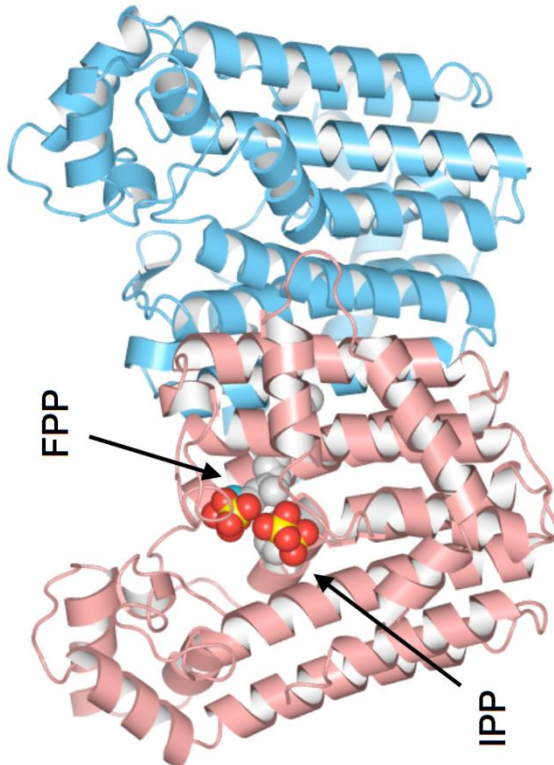
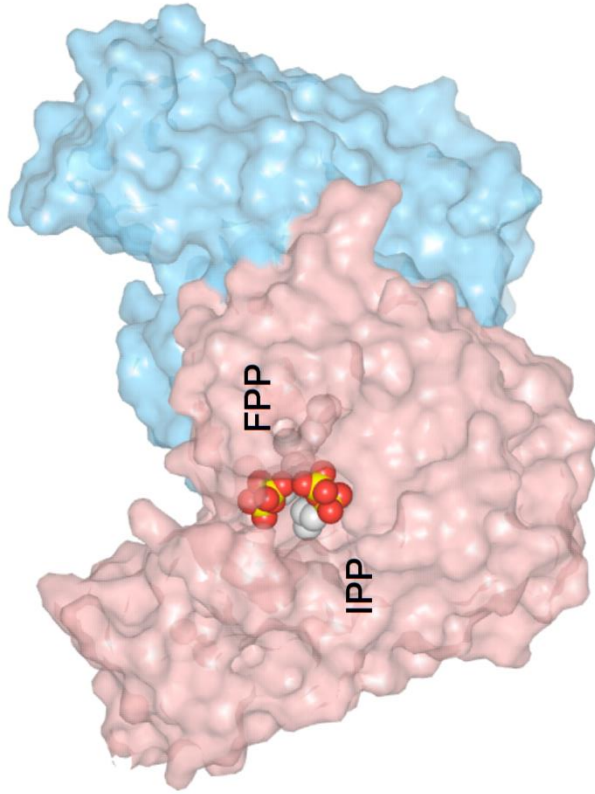


Fig. 23

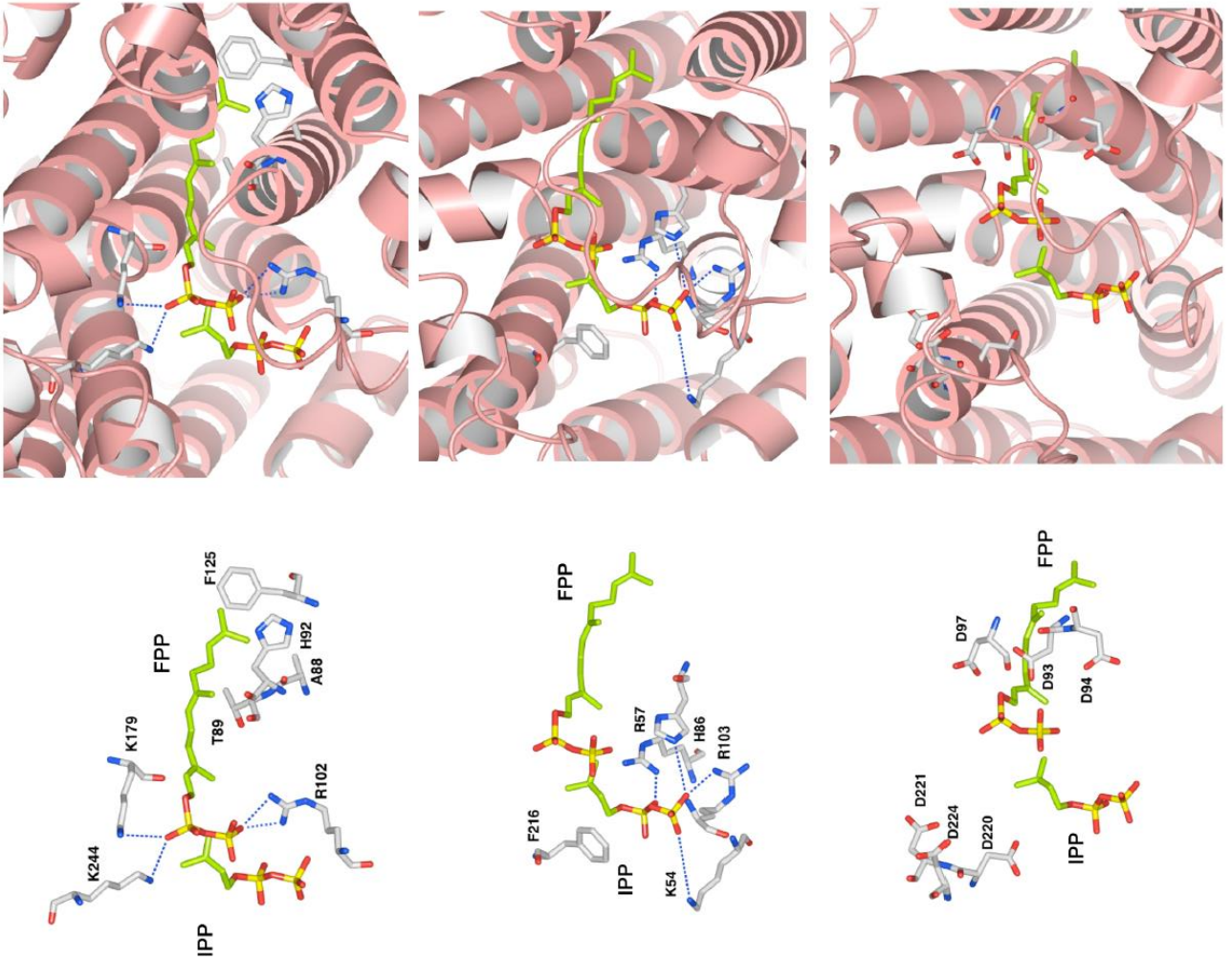
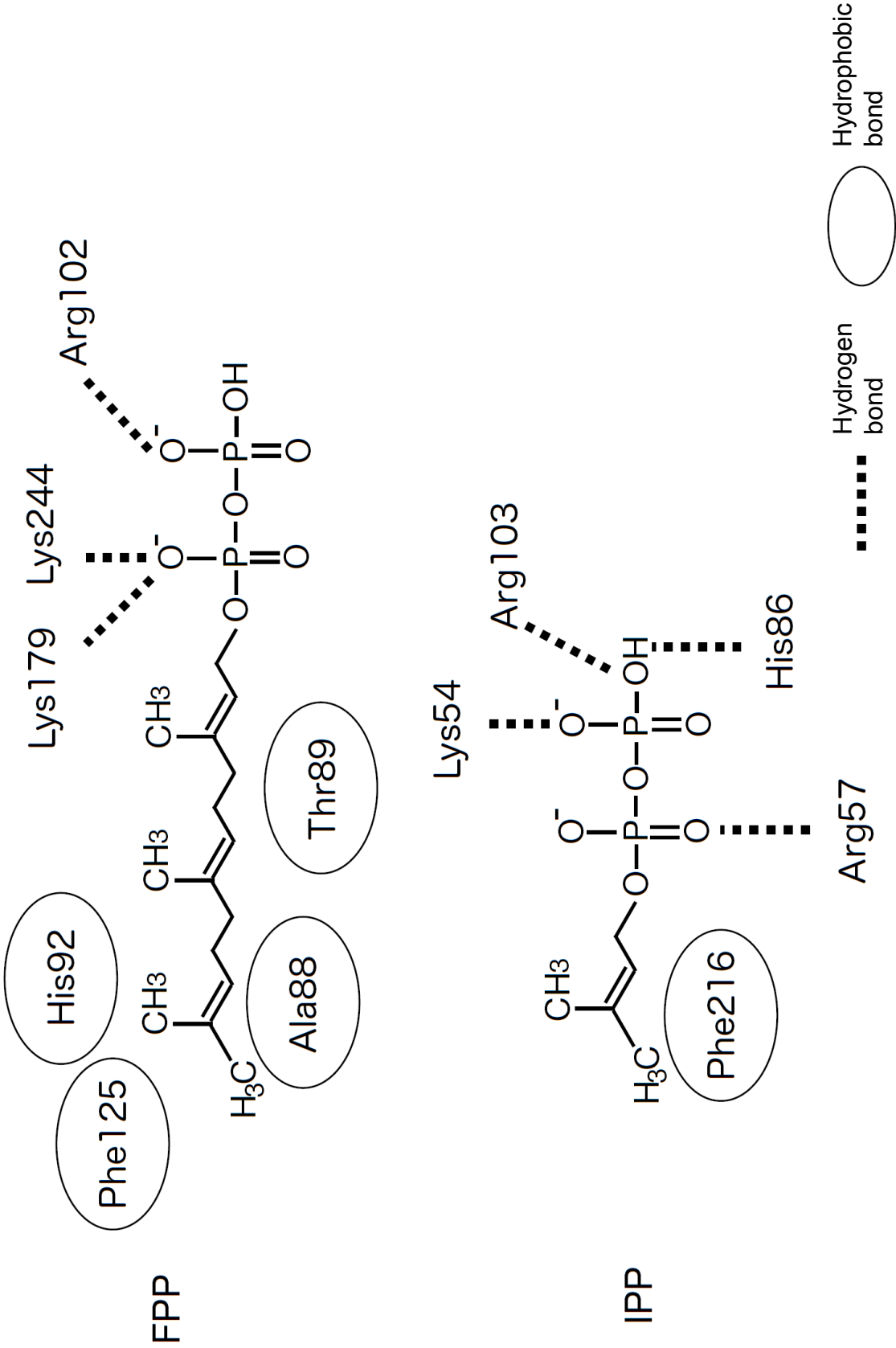


Fig. 24



Asp94

.....

Asp93 ..... Mg ..... Asp97

Asp221

.....

Asp220 ..... Mg ..... Asp224

Fig. 25

Fig. 26

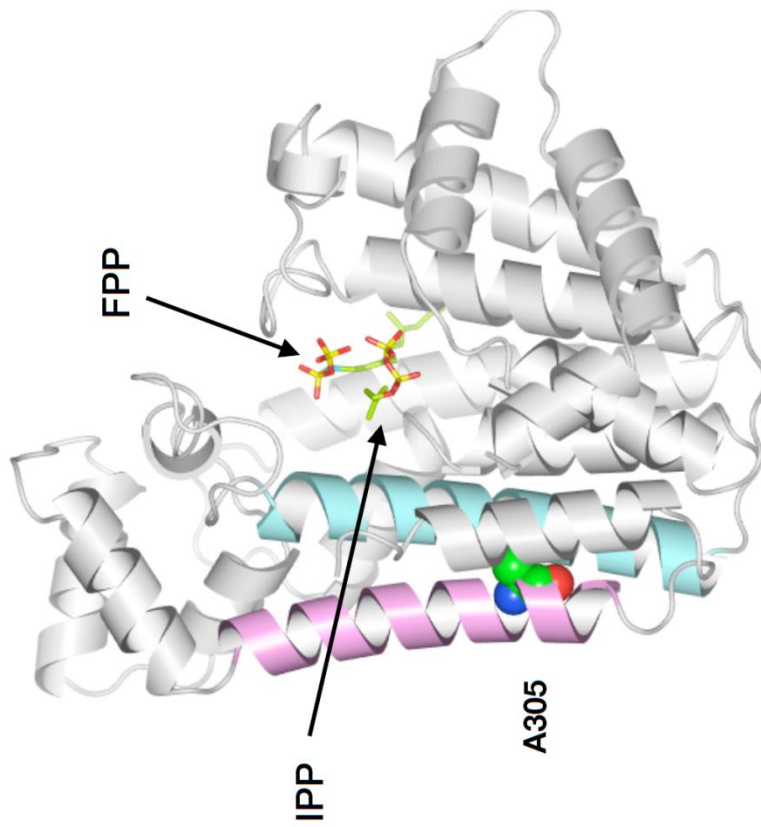
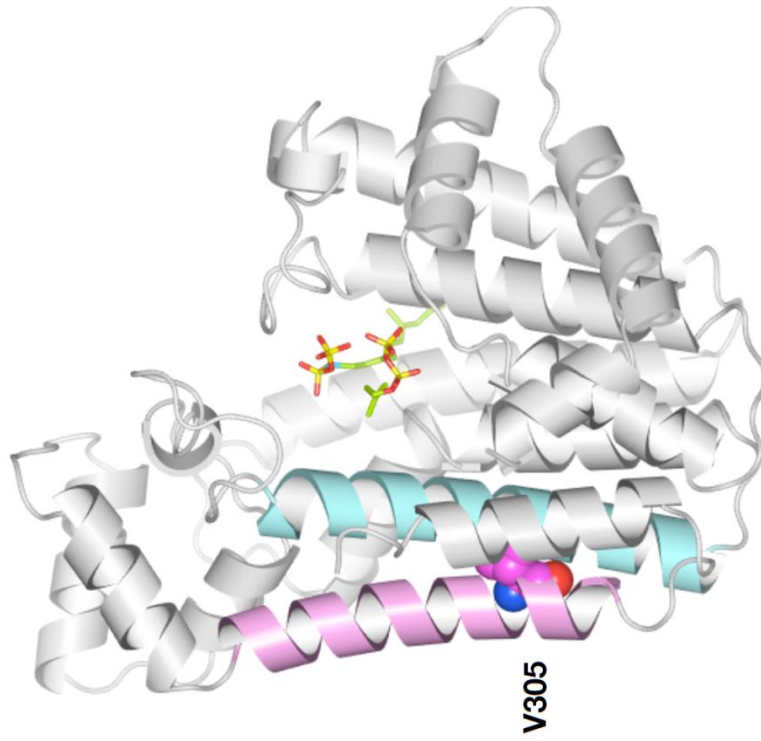


Fig. 27

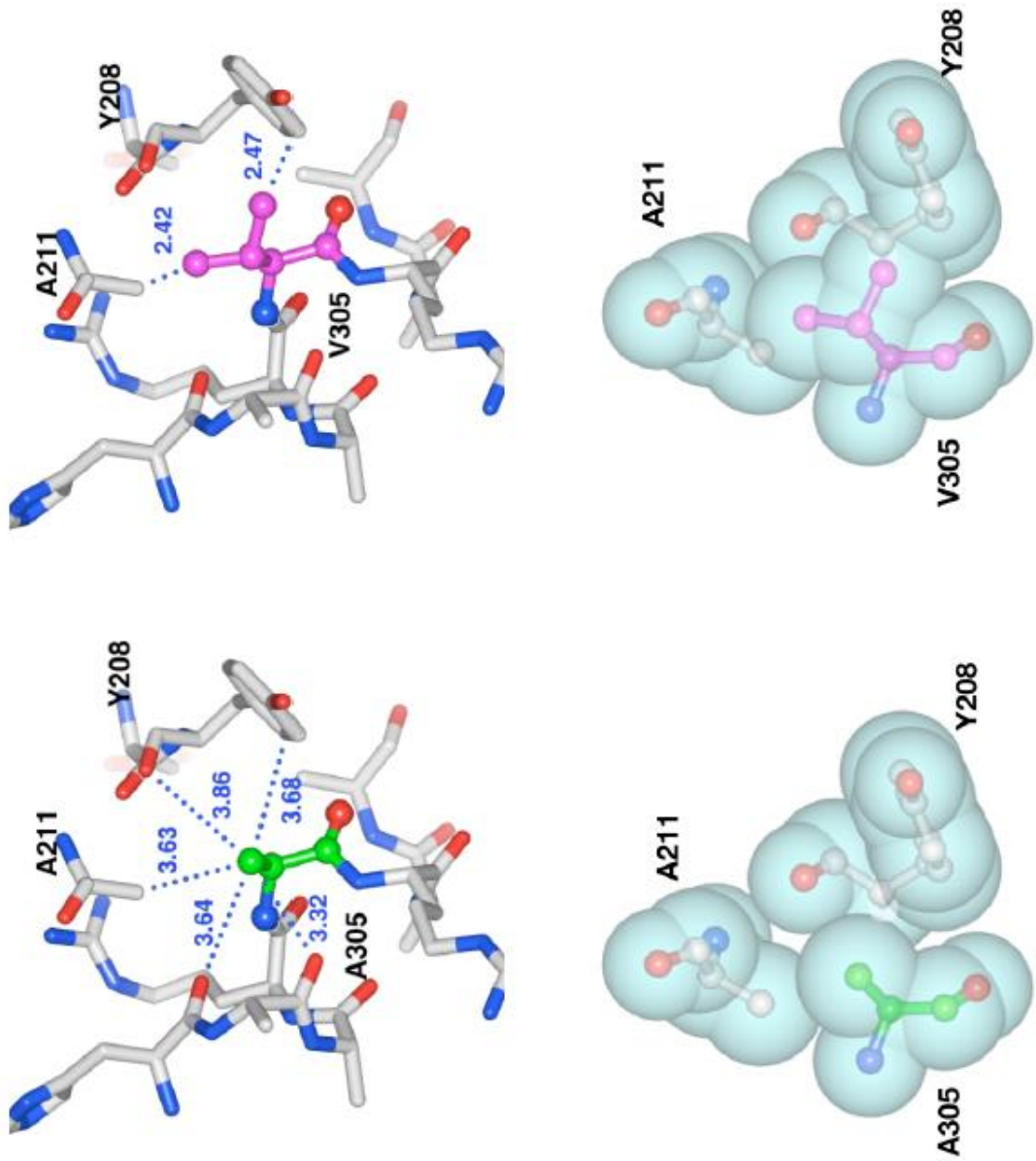


Fig. 28

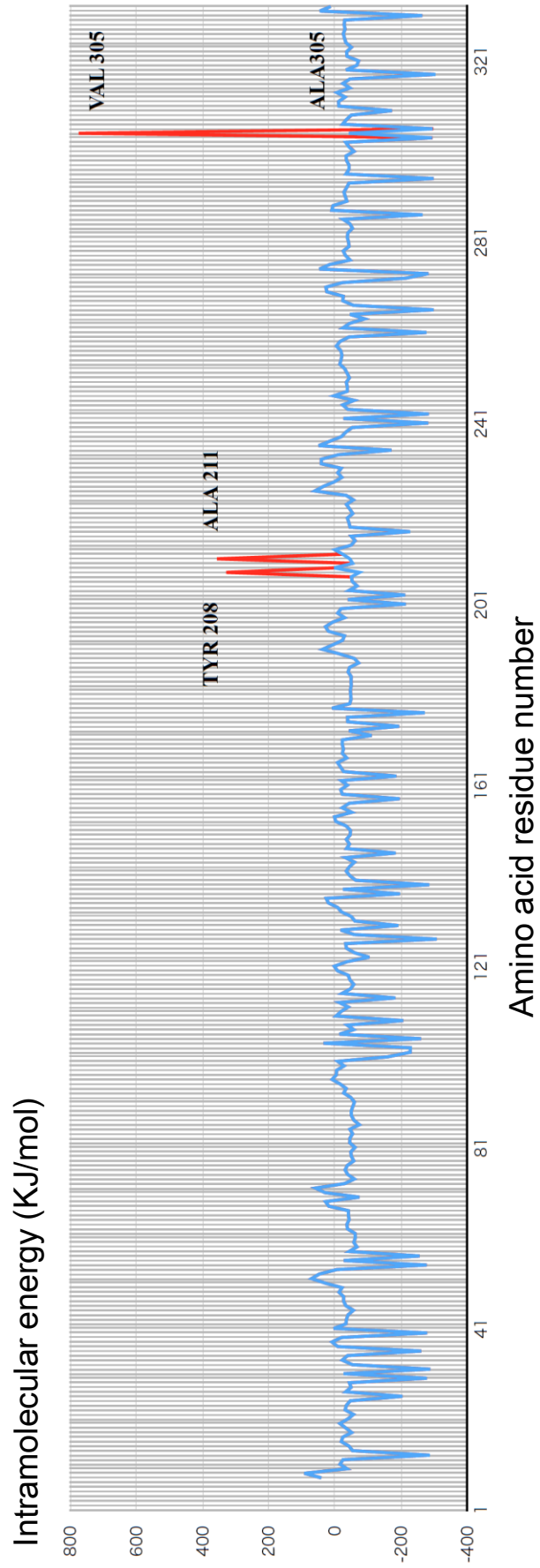




Fig. 29

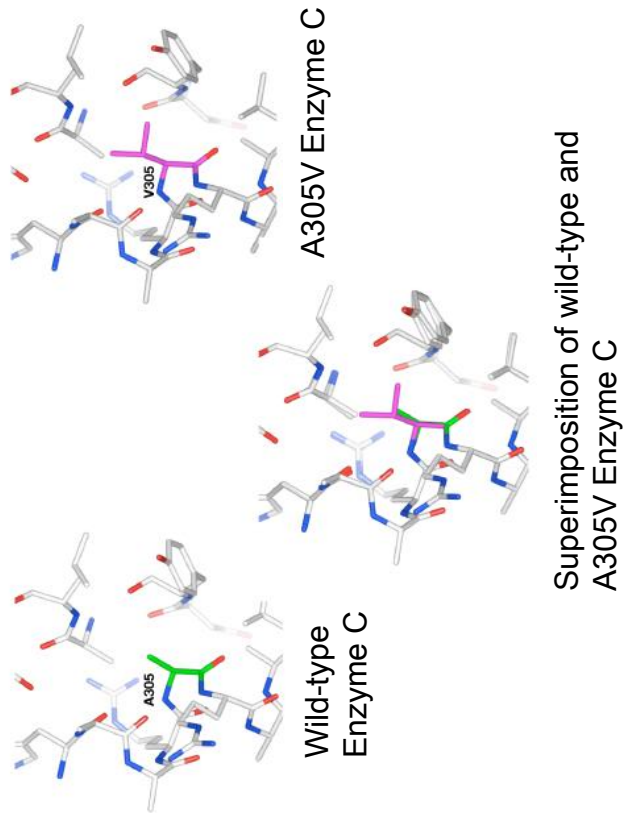
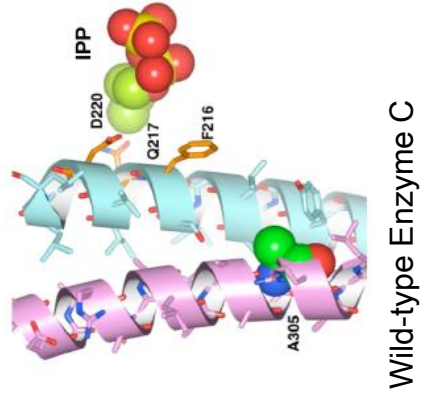
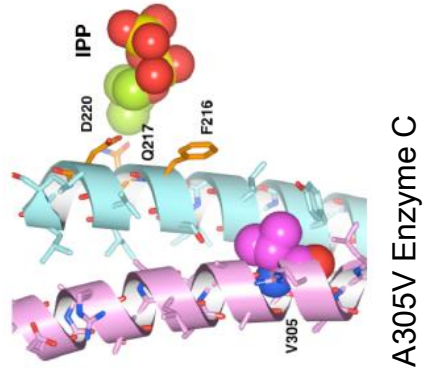




Fig. 30

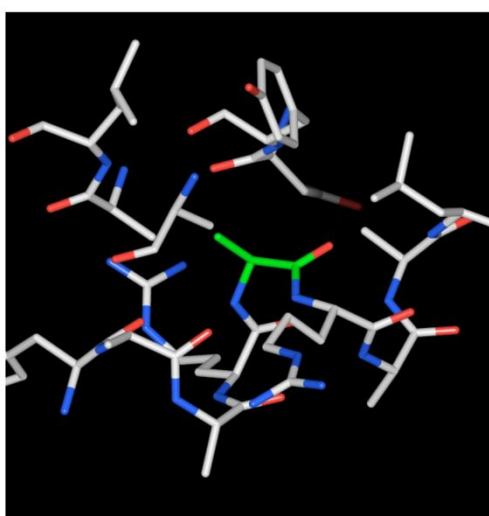
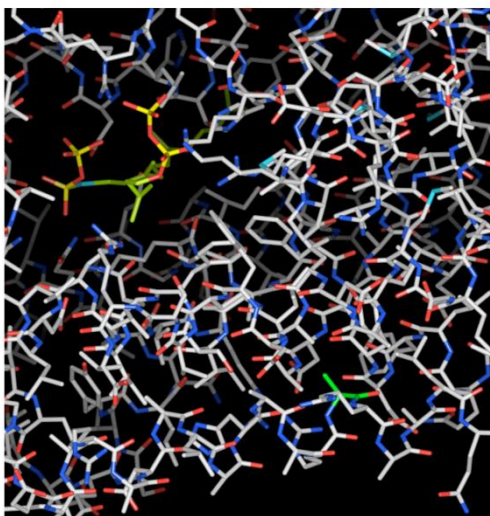
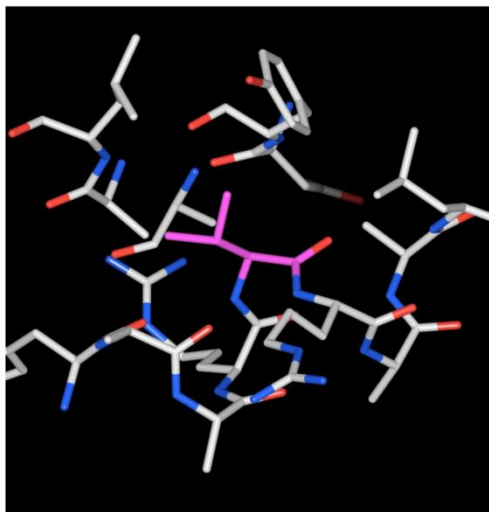
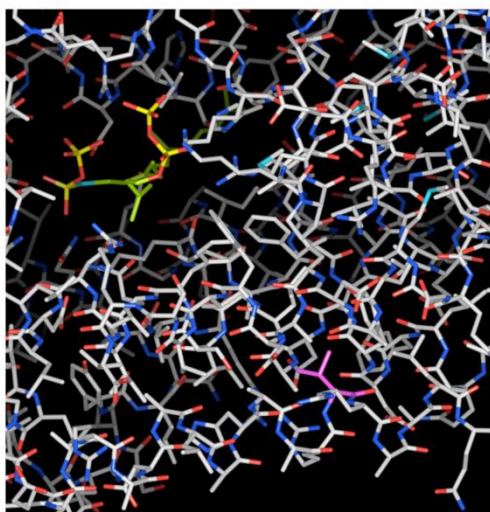
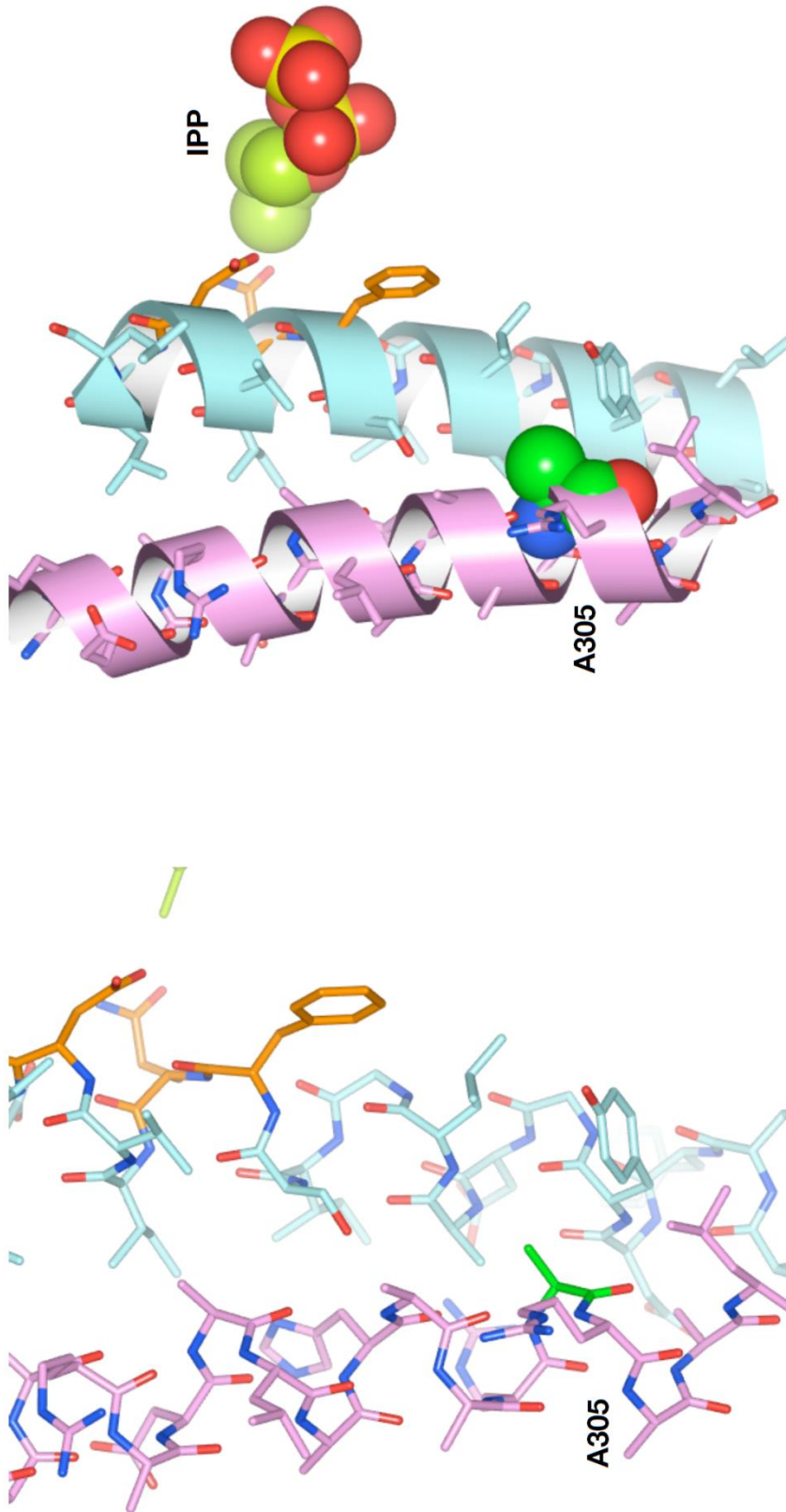
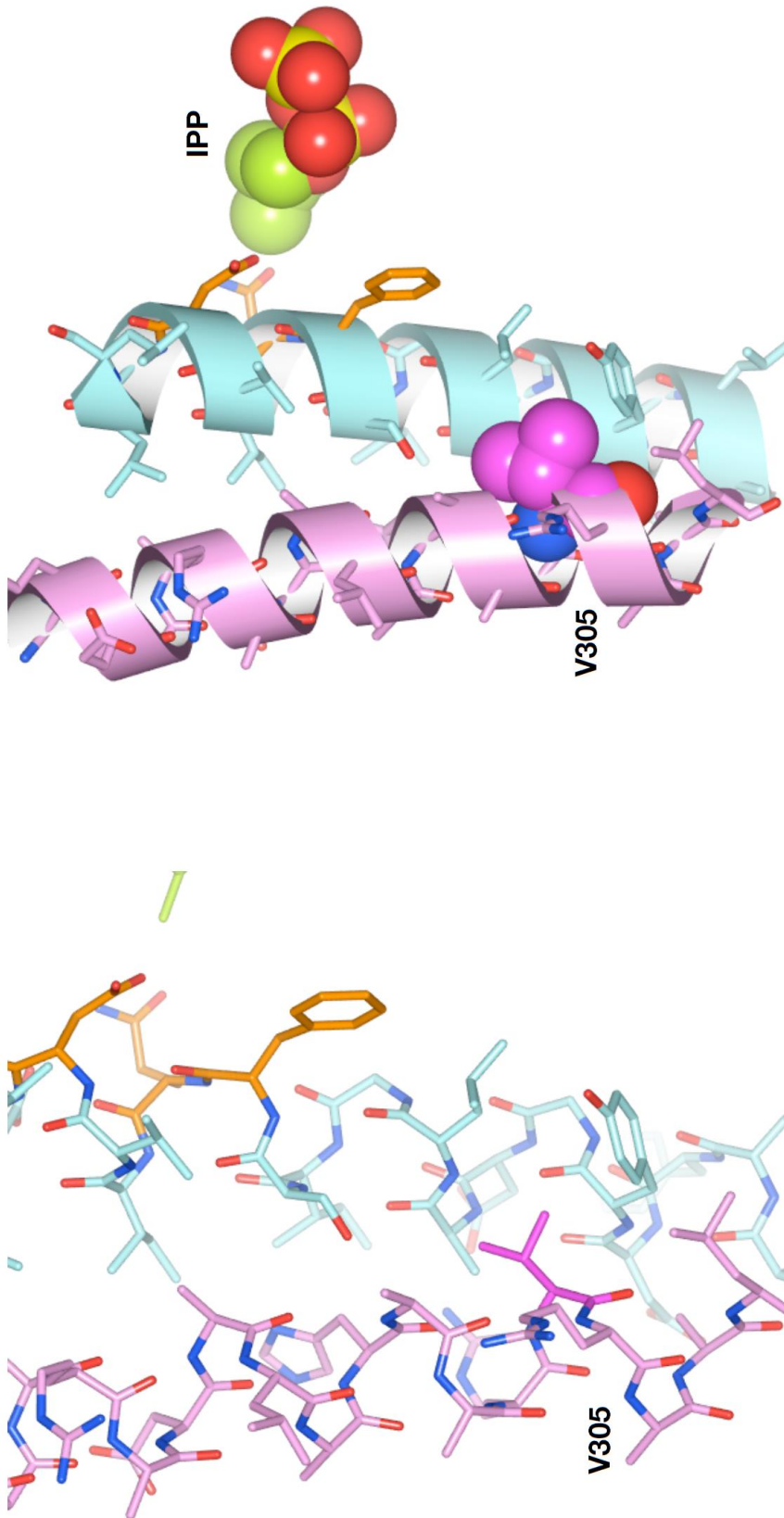


Fig. 31

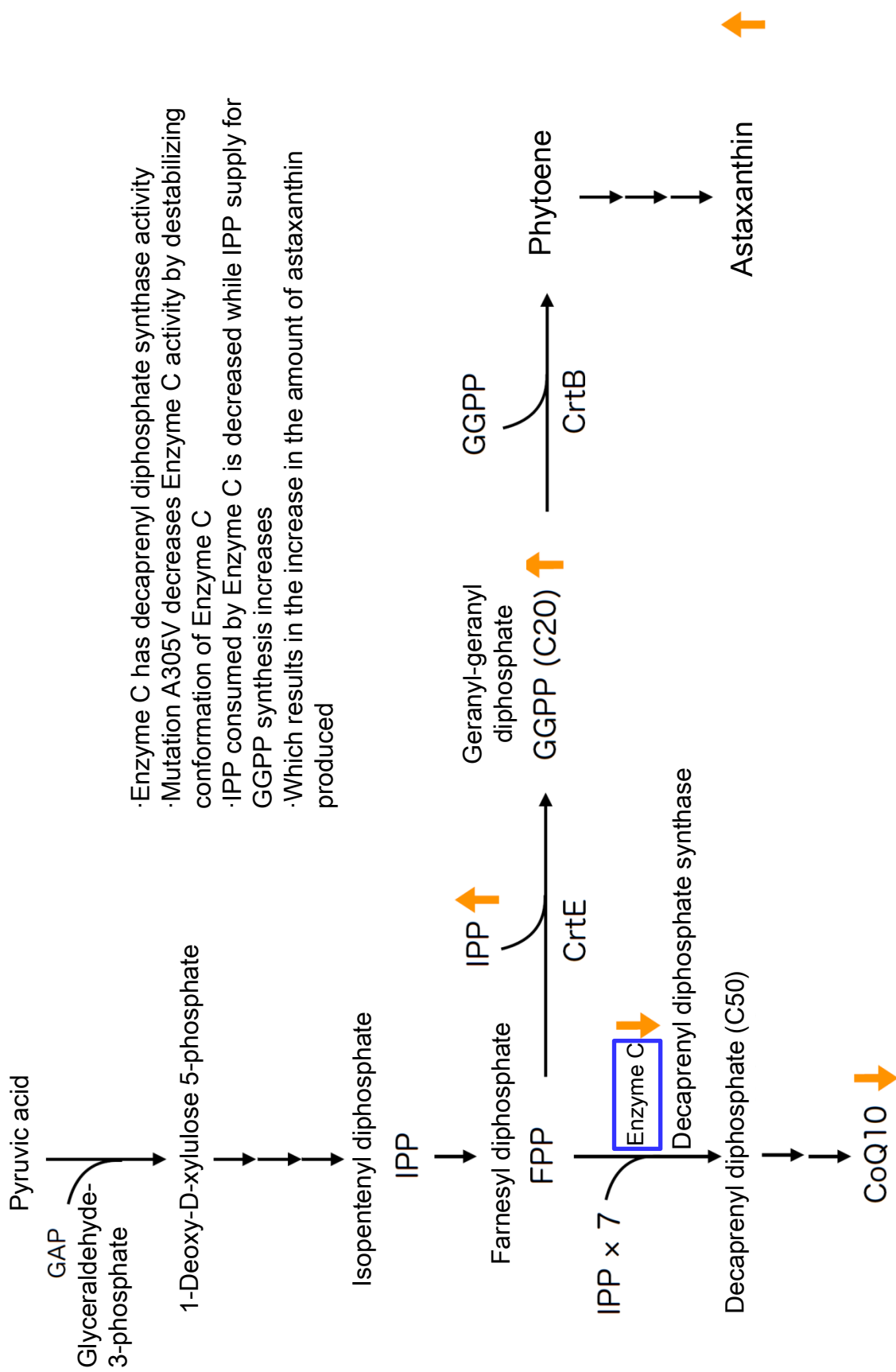


Wild-type Enzyme C

Fig. 32



A305V Enzyme C



- Enzyme C has decaprenyl diphosphate synthase activity
- Mutation A305V decreases Enzyme C activity by destabilizing conformation of Enzyme C
- IPP consumed by Enzyme C is decreased while IPP supply for GGPP synthesis increases
- Which results in the increase in the amount of astaxanthin produced

G1152W0\_Sequence\_ST25.txt  
SEQUENCE LISTING

<110> JXTG NIPPON OIL & ENERGY CORPORATION

<120> A method of producing carotenoids

<130> G1152W0

<150> JP2016-71303

<151> 2016-03-31

<160> 9

<170> PatentIn version 3.5

<210> 1

<211> 1926

<212> DNA

<213> Paracoccus carotini faciens

<220>

<221> CDS

<222> (1)..(1923)

<400> 1

atg	acc	gat	cga	ccc	aag	acg	ccg	att	ctg	gac	cgc	gtt	agc	ctg	ccg	48
Met	Thr	Asp	Arg	Pro	Lys	Thr	Pro	Ile	Leu	Asp	Arg	Val	Ser	Leu	Pro	
1				5					10					15		

tcg	gat	ctg	aag	acc	ctg	acc	gac	gcg	caa	ctg	cgc	cag	ctt	gcc	gac	96
Ser	Asp	Leu	Lys	Thr	Leu	Thr	Asp	Ala	Gln	Leu	Arg	Gln	Leu	Ala	Asp	
			20					25					30			

gag	ttg	cgg	gcc	gag	acc	atc	agc	gcc	gtc	agc	gtg	acc	ggc	ggc	cat	144
Glu	Leu	Arg	Ala	Glu	Thr	Ile	Ser	Ala	Val	Ser	Val	Thr	Gly	Gly	His	
		35					40					45				

ctg	ggc	gcg	ggc	ctg	ggc	gtg	gtc	gag	ctg	acc	gta	gcc	ctg	cat	gcg	192
Leu	Gly	Ala	Gly	Leu	Gly	Val	Val	Glu	Leu	Thr	Val	Ala	Leu	His	Ala	
	50					55					60					

gtc	ttc	gac	acg	ccg	cgc	gac	aag	ctg	atc	tgg	gac	gtg	ggg	cac	cag	240
Val	Phe	Asp	Thr	Pro	Arg	Asp	Lys	Leu	Ile	Trp	Asp	Val	Gly	His	Gln	
65					70					75					80	

tgc	tat	ccg	cac	aag	atc	ctg	acc	ggg	cgg	cgc	gat	cgc	atc	cgc	acg	288
Cys	Tyr	Pro	His	Lys	Ile	Leu	Thr	Gly	Arg	Arg	Asp	Arg	Ile	Arg	Thr	
				85				90						95		

ctg	cgc	atg	ggc	ggg	ggc	ctt	gcg	ggg	ttc	acc	aag	cgc	gcg	gaa	agc	336
Leu	Arg	Met	Gly	Gly	Gly	Leu	Ala	Gly	Phe	Thr	Lys	Arg	Ala	Glu	Ser	
			100					105					110			

ccc	tat	gac	ccg	ttc	ggg	gcg	ggg	cat	tcg	tcg	acc	tcg	atc	agc	gcg	384
Pro	Tyr	Asp	Pro	Phe	Gly	Ala	Gly	His	Ser	Ser	Thr	Ser	Ile	Ser	Ala	
		115					120					125				

gcc	ctg	ggt	ttc	gcg	atg	gcg	cgg	gaa	ctg	ggc	ggc	gac	ccg	ggc	gat	432
Ala	Leu	Gly	Phe	Ala	Met	Ala	Arg	Glu	Leu	Gly	Gly	Asp	Pro	Gly	Asp	
	130					135					140					

gcg	atc	gcg	gtc	atc	ggc	gac	ggc	gcc	atg	agc	gcc	ggc	atg	gct	tat	480
Ala	Ile	Ala	Val	Ile	Gly	Asp	Gly	Ala	Met	Ser	Ala	Gly	Met	Ala	Tyr	
145					150					155					160	

gaa	gcg	ctg	aac	aat	gcg	ggc	cat	gag	ggc	aag	cgc	ctg	ttc	gtg	gtc	528
Glu	Ala	Leu	Asn	Asn	Ala	Gly	His	Glu	Gly	Lys	Arg	Leu	Phe	Val	Val	

## G1152WO\_Sequence\_ST25. txt

															165											170											175	
ctg	aat	gac	aac	gag	atg	tcc	att	gcg	ccc	ccg	gtc	ggc	gcc	atg	tcc	576																						
Leu	Asn	Asp	Asn	Glu	Met	Ser	Ile	Ala	Pro	Pro	Val	Gly	Ala	Met	Ser																							
															180											185											190	
tcc	tat	ctg	acg	cgc	ctt	tat	gca	ggc	ggg	ccg	ttc	cag	gag	ctg	aag	624																						
Ser	Tyr	Leu	Thr	Arg	Leu	Tyr	Ala	Gly	Gly	Pro	Phe	Gln	Glu	Leu	Lys																							
															195											200											205	
gcg	gtg	gcc	aag	ggc	gcc	gtt	ggc	atg	ctg	ccg	gac	gcc	ctg	cag	gag	672																						
Ala	Val	Ala	Lys	Gly	Ala	Val	Gly	Met	Leu	Pro	Asp	Ala	Leu	Gln	Glu																							
															210											215											220	
ggc	gcg	cgc	cgc	gcc	aag	gag	atg	ctg	aag	ggc	atg	gcc	gtg	ggc	ggc	720																						
Gly	Ala	Arg	Arg	Ala	Lys	Glu	Met	Leu	Lys	Gly	Met	Ala	Val	Gly	Gly																							
															225											230											235	
acc	ctg	ttc	gag	gag	ctg	ggc	ttt	tcc	tat	atc	ggg	ccg	gtc	gac	ggt	768																						
Thr	Leu	Phe	Glu	Glu	Leu	Gly	Phe	Ser	Tyr	Ile	Gly	Pro	Val	Asp	Gly																							
															245											250											255	
cac	gac	atg	gaa	cag	ctg	ctg	ccc	ctg	ctg	cgc	acc	gtg	cgg	gcg	cgg	816																						
His	Asp	Met	Glu	Gln	Leu	Leu	Pro	Leu	Leu	Arg	Thr	Val	Arg	Ala	Arg																							
															260											265											270	
gcc	gac	ggt	ccg	gtc	ctg	atc	cac	gtc	gtg	acc	aag	aag	ggc	aag	ggt	864																						
Ala	Asp	Gly	Pro	Val	Leu	Ile	His	Val	Val	Thr	Lys	Lys	Gly	Lys	Gly																							
															275											280											285	
tat	gcc	ccg	gcc	gag	gcg	gcc	gcg	gac	aag	ggc	cat	gcc	acc	gcc	aag	912																						
Tyr	Ala	Pro	Ala	Glu	Ala	Ala	Ala	Asp	Lys	Gly	His	Ala	Thr	Ala	Lys																							
															290											295											300	
ttc	gac	gtg	atc	acc	ggc	gtg	cag	gcc	aag	gcg	aaa	tcg	aac	gcg	ccc	960																						
Phe	Asp	Val	Ile	Thr	Gly	Val	Gln	Ala	Lys	Ala	Lys	Ser	Asn	Ala	Pro																							
															305											310											315	
agc	tat	acc	gcc	gtc	ttc	gcg	cag	gcg	ctg	atc	gat	cag	gcg	ggg	cgt	1008																						
Ser	Tyr	Thr	Ala	Val	Phe	Ala	Gln	Ala	Leu	Ile	Asp	Gln	Ala	Gly	Arg																							
															325											330											335	
gac	gac	cgc	atc	ctg	ggg	gtg	acc	gcc	gcg	atg	ccg	gac	ggg	acc	ggc	1056																						
Asp	Asp	Arg	Ile	Leu	Gly	Val	Thr	Ala	Ala	Met	Pro	Asp	Gly	Thr	Gly																							
															340											345											350	
ctc	aag	cag	ttc	gcg	cag	cgc	ttt	ccg	cgc	cgc	tgc	ttc	gac	gtg	ggc	1104																						
Leu	Lys	Gln	Phe	Ala	Gln	Arg	Phe	Pro	Arg	Arg	Cys	Phe	Asp	Val	Gly																							
															355											360											365	
att	gcc	gaa	cag	cat	gcc	gtg	acc	ttt	gcc	gcc	ggt	ctg	gcc	gcg	gga	1152																						
Ile	Ala	Glu	Gln	His	Ala	Val	Thr	Phe	Ala	Ala	Gly	Leu	Ala	Ala	Gly																							
															370											375											380	
ggc	atg	aag	cct	ttc	gtc	gcg	atc	tat	tcg	acc	ttt	ctg	cag	cgc	ggc	1200																						
Gly	Met	Lys	Pro	Phe	Val	Ala	Ile	Tyr	Ser	Thr	Phe	Leu	Gln	Arg	Gly																							
															385											390											395	
tat	gac	cag	atc	gtg	cat	gac	gtc	gcg	gtc	cag	aac	ctg	ccc	gtc	cgg	1248																						
Tyr	Asp	Gln	Ile	Val	His	Asp	Val	Ala	Val	Gln	Asn	Leu	Pro	Val	Arg																							
															405											410											415	
ttc	gcc	gtc	gac	cgt	gcc	ggg	ctg	gtg	ggc	cag	gac	ggt	ccg	acc	cat	1296																						
Phe	Ala	Val	Asp	Arg	Ala	Gly	Leu	Val	Gly	Gln	Asp	Gly	Pro	Thr	His																							
															420											425											430	
tcc	ggc	gcc	tat	gac	acc	gcc	ttc	ctg	gcc	aat	ctg	ccc	ggg	ttt	gtc	1344																						
Ser	Gly	Ala	Tyr	Asp	Thr	Ala	Phe	Leu	Ala	Asn	Leu	Pro	Gly	Phe	Val																							

G1152W0\_Sequence\_ST25. txt

435	440	445	
gtc atg gcc gcc gcg gac gag gcc gaa ctg gcc cac atg gtc gcc acg Val Met Ala Ala Ala Asp Glu Ala Glu Leu Ala His Met Val Ala Thr 450 455 460			1392
gcc gcc gca cat gac agc ggc ccc atc gcc ttc cgc ttt ccg cgc gga Ala Ala Ala His Asp Ser Gly Pro Ile Ala Phe Arg Phe Pro Arg Gly 465 470 475 480			1440
gag ggc acc ggg gtc gag atg ccc gag cgc ggg cag gtc ctg ccc atc Glu Gly Thr Gly Val Glu Met Pro Glu Arg Gly Gln Val Leu Pro Ile 485 490 495			1488
ggc aag ggc cgc atg atc gcg gag ggc aag ggc gtc gcc atc ctg tcc Gly Lys Gly Arg Met Ile Ala Glu Gly Lys Gly Val Ala Ile Leu Ser 500 505 510			1536
ttc ggc acg cgc ctg tcc gag gtg atg acc gcg cgc gag gcg ctg atg Phe Gly Thr Arg Leu Ser Glu Val Met Thr Ala Arg Glu Ala Leu Met 515 520 525			1584
gcg cgg ggc atc acg ccc acg gtc gcc gac gcg cgc ttt gcc aag ccg Ala Arg Gly Ile Thr Pro Thr Val Ala Asp Ala Arg Phe Ala Lys Pro 530 535 540			1632
ctg gat cgt gac ctg atc ctg gca ctg gcc cgc gat cac gag gcg ctg Leu Asp Arg Asp Leu Ile Leu Ala Leu Ala Arg Asp His Glu Ala Leu 545 550 555 560			1680
atc acc atc gag gag ggc gcc gtg ggc ggt ttc ggc agc cat gtc gcc Ile Thr Ile Glu Glu Gly Ala Val Gly Gly Phe Gly Ser His Val Ala 565 570 575			1728
cat ctg ctg gcc gag gag ggc gcc ttc gac aag ggc ctg cgc ttc cgg His Leu Leu Ala Glu Glu Gly Ala Phe Asp Lys Gly Leu Arg Phe Arg 580 585 590			1776
tcg atg gtc ctg ccg gat gcg ttc gtc gac cac gac gcg ccg cgg gcg Ser Met Val Leu Pro Asp Ala Phe Val Asp His Asp Ala Pro Arg Ala 595 600 605			1824
atg tat gac agc agc tgc ctg aac gcc cag cat atc gag gac aag gtc Met Tyr Asp Ser Ser Cys Leu Asn Ala Gln His Ile Glu Asp Lys Val 610 615 620			1872
ctc tcc gtg atg ggc gtc cag tcg ctc gcg gcg cgg cgc acc gac cgc Leu Ser Val Met Gly Val Gln Ser Leu Ala Ala Arg Arg Thr Asp Arg 625 630 635 640			1920
gcc tga Ala			1926

<210> 2  
 <211> 641  
 <212> PRT  
 <213> Paracoccus carotini faciens

<400> 2

Met Thr Asp Arg Pro Lys Thr Pro Ile Leu Asp Arg Val Ser Leu Pro  
 1 5 10 15

Ser Asp Leu Lys Thr Leu Thr Asp Ala Gln Leu Arg Gln Leu Ala Asp  
 20 25 30

G1152W0\_Sequence\_ST25. txt

Gl u Leu Arg Al a Gl u Thr I l e Ser Al a Val Ser Val Thr Gl y Gl y Hi s  
 35 40 45  
 Leu Gl y Al a Gl y Leu Gl y Val Val Gl u Leu Thr Val Al a Leu Hi s Al a  
 50 55 60  
 Val Phe Asp Thr Pro Arg Asp Lys Leu I l e Trp Asp Val Gl y Hi s Gl n  
 65 70 75 80  
 Cys Tyr Pro Hi s Lys I l e Leu Thr Gl y Arg Arg Asp Arg I l e Arg Thr  
 85 90 95  
 Leu Arg Met Gl y Gl y Gl y Leu Al a Gl y Phe Thr Lys Arg Al a Gl u Ser  
 100 105 110  
 Pro Tyr Asp Pro Phe Gl y Al a Gl y Hi s Ser Ser Thr Ser I l e Ser Al a  
 115 120 125  
 Al a Leu Gl y Phe Al a Met Al a Arg Gl u Leu Gl y Gl y Asp Pro Gl y Asp  
 130 135 140  
 Al a I l e Al a Val I l e Gl y Asp Gl y Al a Met Ser Al a Gl y Met Al a Tyr  
 145 150 155 160  
 Gl u Al a Leu Asn Asn Al a Gl y Hi s Gl u Gl y Lys Arg Leu Phe Val Val  
 165 170 175  
 Leu Asn Asp Asn Gl u Met Ser I l e Al a Pro Pro Val Gl y Al a Met Ser  
 180 185 190  
 Ser Tyr Leu Thr Arg Leu Tyr Al a Gl y Gl y Pro Phe Gl n Gl u Leu Lys  
 195 200 205  
 Al a Val Al a Lys Gl y Al a Val Gl y Met Leu Pro Asp Al a Leu Gl n Gl u  
 210 215 220  
 Gl y Al a Arg Arg Al a Lys Gl u Met Leu Lys Gl y Met Al a Val Gl y Gl y  
 225 230 235 240  
 Thr Leu Phe Gl u Gl u Leu Gl y Phe Ser Tyr I l e Gl y Pro Val Asp Gl y  
 245 250 255  
 Hi s Asp Met Gl u Gl n Leu Leu Pro Leu Leu Arg Thr Val Arg Al a Arg  
 260 265 270  
 Al a Asp Gl y Pro Val Leu I l e Hi s Val Val Thr Lys Lys Gl y Lys Gl y  
 275 280 285  
 Tyr Al a Pro Al a Gl u Al a Al a Al a Asp Lys Gl y Hi s Al a Thr Al a Lys  
 290 295 300



G1152W0\_Sequence\_ST25. txt

Phe Asp Val Ile Thr Gly Val Gln Ala Lys Ala Lys Ser Asn Ala Pro  
 305 310 315 320  
 Ser Tyr Thr Ala Val Phe Ala Gln Ala Leu Ile Asp Gln Ala Gly Arg  
 325 330 335  
 Asp Asp Arg Ile Leu Gly Val Thr Ala Ala Met Pro Asp Gly Thr Gly  
 340 345 350  
 Leu Lys Gln Phe Ala Gln Arg Phe Pro Arg Arg Cys Phe Asp Val Gly  
 355 360 365  
 Ile Ala Glu Gln His Ala Val Thr Phe Ala Ala Gly Leu Ala Ala Gly  
 370 375 380  
 Gly Met Lys Pro Phe Val Ala Ile Tyr Ser Thr Phe Leu Gln Arg Gly  
 385 390 395 400  
 Tyr Asp Gln Ile Val His Asp Val Ala Val Gln Asn Leu Pro Val Arg  
 405 410 415  
 Phe Ala Val Asp Arg Ala Gly Leu Val Gly Gln Asp Gly Pro Thr His  
 420 425 430  
 Ser Gly Ala Tyr Asp Thr Ala Phe Leu Ala Asn Leu Pro Gly Phe Val  
 435 440 445  
 Val Met Ala Ala Ala Asp Glu Ala Glu Leu Ala His Met Val Ala Thr  
 450 455 460  
 Ala Ala Ala His Asp Ser Gly Pro Ile Ala Phe Arg Phe Pro Arg Gly  
 465 470 475 480  
 Glu Gly Thr Gly Val Glu Met Pro Glu Arg Gly Gln Val Leu Pro Ile  
 485 490 495  
 Gly Lys Gly Arg Met Ile Ala Glu Gly Lys Gly Val Ala Ile Leu Ser  
 500 505 510  
 Phe Gly Thr Arg Leu Ser Glu Val Met Thr Ala Arg Glu Ala Leu Met  
 515 520 525  
 Ala Arg Gly Ile Thr Pro Thr Val Ala Asp Ala Arg Phe Ala Lys Pro  
 530 535 540  
 Leu Asp Arg Asp Leu Ile Leu Ala Leu Ala Arg Asp His Glu Ala Leu  
 545 550 555 560  
 Ile Thr Ile Glu Glu Gly Ala Val Gly Gly Phe Gly Ser His Val Ala  
 565 570 575

G1152W0\_Sequence\_ST25. txt

His Leu Leu Ala Glu Glu Gly Ala Phe Asp Lys Gly Leu Arg Phe Arg  
580 585 590

Ser Met Val Leu Pro Asp Ala Phe Val Asp His Asp Ala Pro Arg Ala  
595 600 605

Met Tyr Asp Ser Ser Cys Leu Asn Ala Gln His Ile Glu Asp Lys Val  
610 615 620

Leu Ser Val Met Gly Val Gln Ser Leu Ala Ala Arg Arg Thr Asp Arg  
625 630 635 640

Ala

<210> 3  
<211> 1002  
<212> DNA  
<213> Paracoccus carotini faciens

<220>  
<221> CDS  
<222> (1).. (999)

<400> 3	
atg act gtg cag gac aac gtc cgc aaa ccg atg gat cgg ctg agc gag	48
Met Thr Val Gln Asp Asn Val Arg Lys Pro Met Asp Arg Leu Ser Glu	
1 5 10 15	
gcc ctg acc gcc gag atg gag gcg gtc aac gcg ctg atc cgc gac cgc	96
Ala Leu Thr Ala Glu Met Glu Ala Val Asn Ala Leu Ile Arg Asp Arg	
20 25 30	
atg gcc agc cgc cat gcc ccc cgc atc ccc gag gtg acg gcg cat ctg	144
Met Ala Ser Arg His Ala Pro Arg Ile Pro Glu Val Thr Ala His Leu	
35 40 45	
atc gag gcc ggc ggc aag cgc ctg cgc ccg atg ctg acg ctg gcc gcg	192
Ile Glu Ala Gly Gly Lys Arg Leu Arg Pro Met Leu Thr Leu Ala Ala	
50 55 60	
gcc aag ctg ctg ggc tat ccc ggc cgc tgg cat gtc cac ctg gcc gcc	240
Ala Lys Leu Leu Gly Tyr Pro Gly Pro Trp His Val His Leu Ala Ala	
65 70 75 80	
acg gtc gaa ttc atc cat acc gcc acg ctg ctg cat gac gac gtg gtc	288
Thr Val Glu Phe Ile His Thr Ala Thr Leu Leu His Asp Asp Val Val	
85 90 95	
gac gaa agc gcg cag cgg cgc ggg cgc ccg acg gcg aac ctg ctg tgg	336
Asp Glu Ser Ala Gln Arg Arg Gly Arg Pro Thr Ala Asn Leu Leu Trp	
100 105 110	
gac aac aag tcc agc gtg ctg gtc gga gat tac ctg ttc gcg cgc agc	384
Asp Asn Lys Ser Ser Val Leu Val Gly Asp Tyr Leu Phe Ala Arg Ser	
115 120 125	
ttt cag ctg atg gtc gaa ccc ggc aac ctg cga acg ctg gag atc ctg	432
Phe Gln Leu Met Val Glu Pro Gly Asn Leu Arg Thr Leu Glu Ile Leu	
130 135 140	

G1152W0\_Sequence\_ST25. txt

gcc	aat	gcc	agc	gcc	acc	atc	gcc	gag	ggc	gag	gtg	ctg	cag	ctg	acc	480
Ala	Asn	Ala	Ser	Ala	Thr	Ile	Ala	Glu	Gly	Glu	Val	Leu	Gln	Leu	Thr	
145					150				155						160	
gcg	gcg	cag	gac	ctg	gcg	acc	gac	gaa	tcc	gtc	tat	ctg	cag	gtg	gtg	528
Ala	Ala	Gln	Asp	Leu	Ala	Thr	Asp	Glu	Ser	Val	Tyr	Leu	Gln	Val	Val	
				165					170					175		
cgc	ggc	aag	acg	gcg	gcg	ctg	ttt	tcc	gcc	gcg	acc	gag	gtg	ggc	ggc	576
Arg	Gly	Lys	Thr	Ala	Ala	Leu	Phe	Ser	Ala	Ala	Thr	Glu	Val	Gly	Gly	
			180					185					190			
gtc	atc	gcc	ggt	gcg	ccc	gac	gat	cag	gtg	cag	gcg	ctg	ttc	gat	tac	624
Val	Ile	Ala	Gly	Ala	Pro	Asp	Asp	Gln	Val	Gln	Ala	Leu	Phe	Asp	Tyr	
		195					200					205				
ggc	gac	gcc	ctg	ggc	atc	agc	ttc	cag	att	gtc	gac	gac	ctg	ctg	gat	672
Gly	Asp	Ala	Leu	Gly	Ile	Ser	Phe	Gln	Ile	Val	Asp	Asp	Leu	Leu	Asp	
	210					215					220					
tac	ggc	ggc	gcg	acc	gag	acg	atc	ggc	aag	aac	gtg	ggc	gac	gat	ttc	720
Tyr	Gly	Gly	Ala	Thr	Glu	Thr	Ile	Gly	Lys	Asn	Val	Gly	Asp	Asp	Phe	
225					230					235					240	
cgc	gaa	cgc	aag	ctg	acc	ctg	ccg	gtc	atc	aag	gcc	atc	gcc	aag	gcc	768
Arg	Glu	Arg	Lys	Leu	Thr	Leu	Pro	Val	Ile	Lys	Ala	Ile	Ala	Lys	Ala	
				245					250					255		
gac	gcc	gag	gaa	cgc	gcc	ttc	tgg	acc	cgc	acc	atc	gag	gcg	ggc	gac	816
Asp	Ala	Glu	Glu	Arg	Ala	Phe	Trp	Thr	Arg	Thr	Ile	Glu	Ala	Gly	Asp	
			260					265					270			
cag	cgc	gac	ggc	gac	ctg	gag	cac	gcg	ctg	tcg	ctg	ctg	gcc	cgt	cac	864
Gln	Arg	Asp	Gly	Asp	Leu	Glu	His	Ala	Leu	Ser	Leu	Leu	Ala	Arg	His	
		275					280					285				
ggt	gcg	atg	gag	gcc	gcg	cgc	gcc	gat	gcg	ctg	gcc	cat	gcc	gca	cgg	912
Gly	Ala	Met	Glu	Ala	Ala	Arg	Ala	Asp	Ala	Leu	Ala	His	Ala	Ala	Arg	
	290					295					300					
gcc	cgc	gcg	gcg	ctg	cag	gtg	ctg	ccc	gcg	cat	ccg	att	cgc	gac	atg	960
Ala	Arg	Ala	Ala	Leu	Gln	Val	Leu	Pro	Ala	His	Pro	Ile	Arg	Asp	Met	
305					310					315					320	
ctg	gcc	gac	ctc	gcg	gat	ttc	gtc	gtc	agc	cgc	gtg	gcc	tga			1002
Leu	Ala	Asp	Leu	Ala	Asp	Phe	Val	Val	Ser	Arg	Val	Ala				
				325					330							

<210> 4  
 <211> 333  
 <212> PRT  
 <213> Paracoccus carotini faciens  
 <400> 4

Met	Thr	Val	Gln	Asp	Asn	Val	Arg	Lys	Pro	Met	Asp	Arg	Leu	Ser	Glu
1				5					10					15	
Ala	Leu	Thr	Ala	Glu	Met	Glu	Ala	Val	Asn	Ala	Leu	Ile	Arg	Asp	Arg
			20					25					30		
Met	Ala	Ser	Arg	His	Ala	Pro	Arg	Ile	Pro	Glu	Val	Thr	Ala	His	Leu
		35					40					45			

## G1152W0\_Sequence\_ST25. txt

I l e G l u A l a G l y G l y L y s A r g L e u A r g P r o M e t L e u T h r L e u A l a A l a  
 50 55 60  
 A l a L y s L e u L e u G l y T y r P r o G l y P r o T r p H i s V a l H i s L e u A l a A l a  
 65 70 75 80  
 T h r V a l G l u P h e I l e H i s T h r A l a T h r L e u L e u H i s A s p A s p V a l V a l  
 85 90 95  
 A s p G l u S e r A l a G l n A r g A r g G l y A r g P r o T h r A l a A s n L e u L e u T r p  
 100 105 110  
 A s p A s n L y s S e r S e r V a l L e u V a l G l y A s p T y r L e u P h e A l a A r g S e r  
 115 120 125  
 P h e G l n L e u M e t V a l G l u P r o G l y A s n L e u A r g T h r L e u G l u I l e L e u  
 130 135 140  
 A l a A s n A l a S e r A l a T h r I l e A l a G l u G l y G l u V a l L e u G l n L e u T h r  
 145 150 155 160  
 A l a A l a G l n A s p L e u A l a T h r A s p G l u S e r V a l T y r L e u G l n V a l V a l  
 165 170 175  
 A r g G l y L y s T h r A l a A l a L e u P h e S e r A l a A l a T h r G l u V a l G l y G l y  
 180 185 190  
 V a l I l e A l a G l y A l a P r o A s p A s p G l n V a l G l n A l a L e u P h e A s p T y r  
 195 200 205  
 G l y A s p A l a L e u G l y I l e S e r P h e G l n I l e V a l A s p A s p L e u L e u A s p  
 210 215 220  
 T y r G l y G l y A l a T h r G l u T h r I l e G l y L y s A s n V a l G l y A s p A s p P h e  
 225 230 235 240  
 A r g G l u A r g L y s L e u T h r L e u P r o V a l I l e L y s A l a I l e A l a L y s A l a  
 245 250 255  
 A s p A l a G l u G l u A r g A l a P h e T r p T h r A r g T h r I l e G l u A l a G l y A s p  
 260 265 270  
 G l n A r g A s p G l y A s p L e u G l u H i s A l a L e u S e r L e u L e u A l a A r g H i s  
 275 280 285  
 G l y A l a M e t G l u A l a A l a A r g A l a A s p A l a L e u A l a H i s A l a A l a A r g  
 290 295 300  
 A l a A r g A l a A l a L e u G l n V a l L e u P r o A l a H i s P r o I l e A r g A s p M e t  
 305 310 315 320

G1152W0\_Sequence\_ST25. txt

Leu Ala Asp Leu Ala Asp Phe Val Val Ser Arg Val Ala  
325 330

<210> 5  
<211> 1926  
<212> DNA  
<213> Paracoccus carotini faciens

<220>  
<221> CDS  
<222> (1).. (1923)

<400> 5  
atg acc gat cga ccc aag acg ccg att ctg gac cgc gtt agc ctg ccg 48  
Met Thr Asp Arg Pro Lys Thr Pro Ile Leu Asp Arg Val Ser Leu Pro  
1 5 10 15  
tcg gat ctg aag acc ctg acc gac gcg caa ctg cgc cag ctt gcc gac 96  
Ser Asp Leu Lys Thr Leu Thr Asp Ala Gln Leu Arg Gln Leu Ala Asp  
20 25 30  
gag ttg cgg gcc gag acc atc agc gcc gtc agc gtg acc ggc ggc cat 144  
Glu Leu Arg Ala Glu Thr Ile Ser Ala Val Ser Val Thr Gly Gly His  
35 40 45  
ctg ggc cgc gcc ctg gcc gtg gtc gag ctg acc gta gcc ctg cat gcg 192  
Leu Gly Ala Gly Leu Gly Val Val Glu Leu Thr Val Ala Leu His Ala  
50 55 60  
gtc ttc gac acg ccg cgc gac aag ctg atc tgg gac gtg ggg cac cag 240  
Val Phe Asp Thr Pro Arg Asp Lys Leu Ile Trp Asp Val Gly His Gln  
65 70 75 80  
tgc tat ccg cac aag atc ctg acc ggg cgg cgc gat cgc atc cgc acg 288  
Cys Tyr Pro His Lys Ile Leu Thr Gly Arg Arg Asp Arg Ile Arg Thr  
85 90 95  
ctg cgc atg gcc ggg gcc ctt gcg ggg ttc acc aag cgc gcc gaa agc 336  
Leu Arg Met Gly Gly Gly Leu Ala Gly Phe Thr Lys Arg Ala Glu Ser  
100 105 110  
ccc tat gac ccg ttc ggg gcg ggg cat tcg tcg acc tcg atc agc gcg 384  
Pro Tyr Asp Pro Phe Gly Ala Gly His Ser Ser Thr Ser Ile Ser Ala  
115 120 125  
gcc ctg ggt ttc gcg atg gcg ccg gaa ctg gcc gcc gac ccg gcc gat 432  
Ala Leu Gly Phe Ala Met Ala Arg Glu Leu Gly Gly Asp Pro Gly Asp  
130 135 140  
gcg atc gcg gtc atc ggc gac gcc ggc atg agc gcc gcc atg gct tat 480  
Ala Ile Ala Val Ile Gly Asp Gly Ala Met Ser Ala Gly Met Ala Tyr  
145 150 155 160  
gaa gcg ctg aac aat gcg gcc cat gag gcc aag cgc ctg ttc gtg gtc 528  
Glu Ala Leu Asn Asn Ala Gly His Glu Gly Lys Arg Leu Phe Val Val  
165 170 175  
ctg aat gac aac gag atg tcc att gcg ccc ccg gtc gcc gcc atg tcc 576  
Leu Asn Asp Asn Glu Met Ser Ile Ala Pro Pro Val Gly Ala Met Ser  
180 185 190  
tcc tat ctg acg cgc ctt tat gca gcc ggg ccg ttc cag gag ctg aag 624  
Ser Tyr Leu Thr Arg Leu Tyr Ala Gly Gly Pro Phe Gln Glu Leu Lys  
195 200 205

## G1152W0\_Sequence\_ST25. txt

gcg Ala 210	gtg Val 210	gcc Ala 210	aag Lys 210	ggc Gly 210	gcc Ala 215	ggt Val 215	ggc Gly 215	atg Met 215	ctg Leu 215	ccg Pro 220	gac Asp 220	gcc Ala 220	ctg Leu 220	cag Gln 220	gag Glu 220	672
gac Asp 225	gcg Ala 225	cgc Arg 225	cgc Arg 225	gcc Ala 230	aag Lys 230	gag Glu 230	atg Met 230	ctg Leu 230	aag Lys 235	ggc Gly 235	atg Met 235	gcc Ala 235	gtg Val 235	ggc Gly 240	ggc Gly 240	720
acc Thr 245	ctg Leu 245	ttc Phe 245	gag Glu 245	gag Glu 245	ctg Leu 245	ggc Gly 245	ttt Phe 250	tcc Ser 250	tat Tyr 250	atc Ile 250	ggg Gly 250	ccg Pro 250	gtc Val 255	gac Asp 255	ggt Gly 255	768
cac His 260	gac Asp 260	atg Met 260	gaa Glu 260	cag Gln 260	ctg Leu 260	ctg Leu 265	ccc Pro 265	ctg Leu 265	ctg Leu 265	cgc Arg 265	acc Thr 270	gtg Val 270	cgg Arg 270	gcg Ala 270	cgg Arg 270	816
gcc Ala 275	gac Asp 275	ggt Gly 275	ccg Pro 275	gtc Val 275	ctg Leu 275	atc Ile 280	cac His 280	gtc Val 280	gtg Val 280	acc Thr 280	aag Lys 285	aag Lys 285	ggc Gly 285	aag Lys 285	ggt Gly 285	864
tat Tyr 290	gcc Ala 290	ccg Pro 290	gcc Ala 290	gag Glu 290	gcg Ala 295	gcc Ala 295	gcg Ala 295	gac Asp 295	aag Lys 295	ggc Gly 295	cat His 300	gcc Ala 300	acc Thr 300	gcc Ala 300	aag Lys 300	912
ttc Phe 305	gac Asp 305	gtg Val 305	atc Ile 305	acc Thr 305	ggc Gly 310	gtg Val 310	cag Gln 310	gcc Ala 310	aag Lys 310	gcg Ala 315	aaa Lys 315	tcg Ser 315	aac Asn 315	gcg Ala 315	ccc Pro 320	960
agc Ser 325	tat Tyr 325	acc Thr 325	gcc Ala 325	gtc Val 325	ttc Phe 325	gcg Ala 325	cag Gln 325	gcg Ala 325	ctg Leu 330	atc Ile 330	gat Asp 330	cag Gln 330	gcg Ala 330	ggg Gly 335	cgt Arg 335	1008
gac Asp 340	gac Asp 340	cgc Arg 340	atc Ile 340	ctg Leu 340	ggg Gly 340	gtg Val 340	acc Thr 340	gcc Ala 345	gcg Ala 345	atg Met 345	ccg Pro 345	gac Asp 345	ggg Gly 350	acc Thr 350	ggc Gly 350	1056
ctc Leu 355	aag Lys 355	cag Gln 355	ttc Phe 355	gcg Ala 355	cag Gln 355	cgc Arg 355	ttt Phe 360	ccg Pro 360	cgc Arg 360	cgc Arg 360	tgc Cys 360	ttc Phe 365	gac Asp 365	gtg Val 365	ggc Gly 365	1104
att Ile 370	gcc Ala 370	gaa Glu 370	cag Gln 370	cat His 370	gcc Ala 370	gtg Val 375	acc Thr 375	ttt Phe 375	gcc Ala 375	gcc Ala 375	ggt Gly 380	ctg Leu 380	gcc Ala 380	gcg Ala 380	gga Gly 380	1152
ggc Gly 385	atg Met 385	aag Lys 385	cct Pro 385	ttc Phe 385	gtc Val 390	gcg Ala 390	atc Ile 390	tat Tyr 390	tcg Ser 395	acc Thr 395	ttt Phe 395	ctg Leu 395	cag Gln 395	cgc Arg 395	ggc Gly 400	1200
tat Tyr 405	gac Asp 405	cag Gln 405	atc Ile 405	gtg Val 405	cat His 405	gac Asp 405	gtc Val 405	gcg Ala 410	gtc Val 410	cag Gln 410	aac Asn 410	ctg Leu 410	ccc Pro 415	gtc Val 415	cgg Arg 415	1248
ttc Phe 420	gcc Ala 420	gtc Val 420	gac Asp 420	cgt Arg 420	gcc Ala 420	ggg Gly 420	ctg Leu 425	gtg Val 425	ggc Gly 425	cag Gln 425	gac Asp 425	ggt Gly 430	ccg Pro 430	acc Thr 430	cat His 430	1296
tcc Ser 435	ggc Gly 435	gcc Ala 435	tat Tyr 435	gac Asp 435	acc Thr 435	gcc Ala 440	ttc Phe 440	ctg Leu 440	gcc Ala 440	aac Asn 440	ctg Leu 445	ccc Pro 445	ggg Gly 445	ttt Phe 445	gtc Val 445	1344
gtc Val 450	atg Met 450	gcc Ala 450	gcc Ala 450	gcg Ala 450	gac Asp 450	gag Glu 455	gcc Ala 455	gaa Glu 455	ctg Leu 455	gcc Ala 455	cac His 460	atg Met 460	gtc Val 460	gcc Ala 460	acg Thr 460	1392
gcc Ala 465	gcc Ala 465	gca Ala 465	cat His 465	gac Asp 465	agc Ser 470	ggc Gly 470	ccc Pro 470	atc Ile 470	gcc Ala 475	ttc Phe 475	cgc Arg 475	ttt Phe 475	ccg Pro 475	cgc Arg 475	gga Gly 480	1440

G1152W0\_Sequence\_ST25. txt

gag ggc acc ggg gtc gag atg ccc gag cgc ggg cag gtc ctg ccc atc Glu Gly Thr Gly Val 485 Glu Met Pro Glu Arg 490 Gly Gl n Val Leu Pro 495 Ile	1488
ggc aag ggc cgc atg atc gcg gag ggc aag ggc gtc gcc atc ctg tcc Gly Lys Gly Arg 500 Met Ile Ala Glu Gly 505 Lys Gly Val Ala Ile Leu Ser	1536
ttc ggc acg cgc ctg tcc gag gtg atg acc gcg cgc gag gcg ctg atg Phe Gly Thr Arg Leu Ser Glu Val 520 Met Thr Ala Arg 525 Glu Ala Leu Met	1584
gcg cgg ggc atc acg ccc acg gtc gcc gac gcg cgc ttt gcc aag ccg Ala Arg 530 Gly Ile Thr Pro 535 Val Ala Asp Ala Arg 540 Phe Ala Lys Pro	1632
ctg gat cgt gac ctg atc ctg gca ctg gcc cgc gat cac gag gcg ctg Leu Asp Arg Asp Leu Ile Leu Ala Leu Ala Arg 555 Asp His Glu Ala Leu 560	1680
atc acc atc gag gag ggc gcc gtg ggc ggt ttc ggc agc cat gtc gcc Ile Thr Ile Glu Glu Gly Ala Val Gly 570 Phe Gly Ser His Val Ala 575	1728
cat ctg ctg gcc gag gag ggc gcc ttc gac aag ggg ctg cgc ttc cgg His Leu Leu Ala 580 Glu Glu Gly Ala Phe 585 Asp Lys Gly Leu Arg 590 Phe Arg	1776
tcg atg gtc ctg ccg gat gcg ttc gtc gac cac gac gcg ccg cgg gcg Ser Met Val Leu Pro Asp Ala Phe Val Asp His Asp Ala Pro Arg Ala 605	1824
atg tat gac agc agc tgc ctg aac gcc cag cat atc gag gac aag gtc Met Tyr Asp Ser Ser Cys Leu Asn Ala Gl n His Ile Glu Asp Lys Val 610	1872
ctc tcc gtg atg ggc gtc cag tcg ctc gcg gcg cgg cgc acc gac cgc Leu Ser Val Met Gly Val 630 Gl n Ser Leu Ala Ala Arg 635 Arg Thr Asp Arg 640	1920
gcc tga Ala	1926

<210> 6  
 <211> 641  
 <212> PRT  
 <213> Paracoccus carotini faciens

<400> 6

Met Thr Asp Arg Pro Lys Thr Pro Ile Leu Asp Arg Val Ser Leu Pro 1 5 10 15
Ser Asp Leu Lys Thr Leu Thr Asp Ala Gl n Leu Arg Gl n Leu Ala Asp 20 25 30
Glu Leu Arg Ala Glu Thr Ile Ser Ala Val Ser Val Thr Gly Gly His 35 40 45
Leu Gly Ala Gly Leu Gly Val Val Glu Leu Thr Val Ala Leu His Ala 50 55 60

## G1152W0\_Sequence\_ST25. txt

Val Phe Asp Thr Pro Arg Asp Lys Leu Ile Trp Asp Val Gly His Gln  
 65 70 75 80  
 Cys Tyr Pro His Lys Ile Leu Thr Gly Arg Arg Asp Arg Ile Arg Thr  
 85 90 95  
 Leu Arg Met Gly Gly Gly Leu Ala Gly Phe Thr Lys Arg Ala Glu Ser  
 100 105 110  
 Pro Tyr Asp Pro Phe Gly Ala Gly His Ser Ser Thr Ser Ile Ser Ala  
 115 120 125  
 Ala Leu Gly Phe Ala Met Ala Arg Glu Leu Gly Gly Asp Pro Gly Asp  
 130 135 140  
 Ala Ile Ala Val Ile Gly Asp Gly Ala Met Ser Ala Gly Met Ala Tyr  
 145 150 155 160  
 Glu Ala Leu Asn Asn Ala Gly His Glu Gly Lys Arg Leu Phe Val Val  
 165 170 175  
 Leu Asn Asp Asn Glu Met Ser Ile Ala Pro Pro Val Gly Ala Met Ser  
 180 185 190  
 Ser Tyr Leu Thr Arg Leu Tyr Ala Gly Gly Pro Phe Gln Glu Leu Lys  
 195 200 205  
 Ala Val Ala Lys Gly Ala Val Gly Met Leu Pro Asp Ala Leu Gln Glu  
 210 215 220  
 Asp Ala Arg Arg Ala Lys Glu Met Leu Lys Gly Met Ala Val Gly Gly  
 225 230 235 240  
 Thr Leu Phe Glu Glu Leu Gly Phe Ser Tyr Ile Gly Pro Val Asp Gly  
 245 250 255  
 His Asp Met Glu Gln Leu Leu Pro Leu Leu Arg Thr Val Arg Ala Arg  
 260 265 270  
 Ala Asp Gly Pro Val Leu Ile His Val Val Thr Lys Lys Gly Lys Gly  
 275 280 285  
 Tyr Ala Pro Ala Glu Ala Ala Ala Asp Lys Gly His Ala Thr Ala Lys  
 290 295 300  
 Phe Asp Val Ile Thr Gly Val Gln Ala Lys Ala Lys Ser Asn Ala Pro  
 305 310 315 320  
 Ser Tyr Thr Ala Val Phe Ala Gln Ala Leu Ile Asp Gln Ala Gly Arg  
 325 330 335



G1152W0\_Sequence\_ST25.txt

Asp Asp Arg Ile Leu Gly Val Thr Ala Ala Met Pro Asp Gly Thr Gly  
340 345 350

Leu Lys Gln Phe Ala Gln Arg Phe Pro Arg Arg Cys Phe Asp Val Gly  
355 360 365

Ile Ala Glu Gln His Ala Val Thr Phe Ala Ala Gly Leu Ala Ala Gly  
370 375 380

Gly Met Lys Pro Phe Val Ala Ile Tyr Ser Thr Phe Leu Gln Arg Gly  
385 390 395 400

Tyr Asp Gln Ile Val His Asp Val Ala Val Gln Asn Leu Pro Val Arg  
405 410 415

Phe Ala Val Asp Arg Ala Gly Leu Val Gly Gln Asp Gly Pro Thr His  
420 425 430

Ser Gly Ala Tyr Asp Thr Ala Phe Leu Ala Asn Leu Pro Gly Phe Val  
435 440 445

Val Met Ala Ala Ala Asp Glu Ala Glu Leu Ala His Met Val Ala Thr  
450 455 460

Ala Ala Ala His Asp Ser Gly Pro Ile Ala Phe Arg Phe Pro Arg Gly  
465 470 475 480

Glu Gly Thr Gly Val Glu Met Pro Glu Arg Gly Gln Val Leu Pro Ile  
485 490 495

Gly Lys Gly Arg Met Ile Ala Glu Gly Lys Gly Val Ala Ile Leu Ser  
500 505 510

Phe Gly Thr Arg Leu Ser Glu Val Met Thr Ala Arg Glu Ala Leu Met  
515 520 525

Ala Arg Gly Ile Thr Pro Thr Val Ala Asp Ala Arg Phe Ala Lys Pro  
530 535 540

Leu Asp Arg Asp Leu Ile Leu Ala Leu Ala Arg Asp His Glu Ala Leu  
545 550 555 560

Ile Thr Ile Glu Glu Gly Ala Val Gly Gly Phe Gly Ser His Val Ala  
565 570 575

His Leu Leu Ala Glu Glu Gly Ala Phe Asp Lys Gly Leu Arg Phe Arg  
580 585 590

Ser Met Val Leu Pro Asp Ala Phe Val Asp His Asp Ala Pro Arg Ala  
595 600 605

Met Tyr Asp Ser Ser Cys Leu Asn Ala Gl n Hi s Tle Gl u Asp Lys Val  
610 615 620

Leu Ser Val Met Gly Val Gln Ser Leu Ala Ala Arg Arg Thr Asp Arg  
625 630 635 640

Al a

<210>	7
<211>	1002
<212>	DNA
<213>	Paracoccus carotini faciens

<220>  
<221> CDS  
<222> (1)..(999)

<400> atg Met 1	act Thr	gtg Val	cag Gl n	gac Asp 5	aac Asn	gtc Val	cgc Arg	aaa Lys	ccg Pro 10	atg Met	gat Asp	cgg Arg	ctg Leu	agc Ser 15	gag Gl u	48
gcc Ala	ctg Leu	acc Thr	gcc Ala 20	gag Gl u	atg Met	gag Gl u	gcg Ala	gtc Val 25	aac Asn	gcg Ala	ctg Leu	atc Ile	cgc Arg 30	gac Asp	cgc Arg	96
atg Met	gcc Ala	agc Ser 35	cgc Arg	cat Hi s	gcc Ala	ccc Pro	cgc Arg 40	atc Ile	ccc Pro	gag Gl u	gtg Val	acg Thr 45	gcg Ala	cat Hi s	ctg Leu	144
atc Ile 50	gag Gl u	gcc Ala	ggc Gly	ggc Gly	aag Lys	cgc Arg 55	ctg Leu	cgc Arg	ccg Pro	atg Met	ctg Leu 60	acg Thr	ctg Leu	gcc Ala	gcg Ala	192
gcc Ala 65	aag Lys	ctg Leu	ctg Leu	ggc Gly	tat Tyr 70	ccc Pro	ggc Gly	ccg Pro	tgg Trp	cat Hi s 75	gtc Val	cac Hi s	ctg Leu	gcc Ala	gcc Ala 80	240
acg Thr	gtc Val	gaa Gl u	ttc Phe	atc Ile 85	cat Hi s	acc Thr	gcc Ala	acg Thr	ctg Leu 90	ctg Leu	cat Hi s	gac Asp	gac Asp	gtg Val 95	gtc Val	288
gac Asp	gaa Gl u	agc Ser	gcg Ala 100	cag Gl n	cgg Arg	cgc Arg	ggg Gly	cgc Arg 105	ccg Pro	acg Thr	gcg Ala	aac Asn	ctg Leu 110	ctg Leu	tgg Trp	336
gac Asp	aac Asn	aag Lys 115	tcc Ser	agc Ser	gtg Val	ctg Leu	gtc Val 120	gga Gly	gat Asp	tac Tyr	ctg Leu	ttc Phe 125	gcg Ala	cgc Arg	agc Ser	384
ttt Phe	cag Gl n 130	ctg Leu	atg Met	gtc Val	gaa Gl u	ccc Pro 135	ggc Gly	aac Asn	ctg Leu	cga Arg	acg Thr 140	ctg Leu	gag Gl u	atc Ile	ctg Leu	432
gcc Ala 145	aat Asn	gcc Ala	agc Ser	gcc Ala	acc Thr 150	atc Ile	gcc Ala	gag Gl u	ggc Gly	gag Gl u 155	gtg Val	ctg Leu	cag Gl n	ctg Leu	acc Thr 160	480
gcg Ala	gcg Ala	cag Gl n	gac Asp	ctg Leu 165	gcg Ala	acc Thr	gac Asp	gaa Gl u	tcc Ser 170	gtc Val	tat Tyr	ctg Leu	cag Gl n	gtg Val 175	gtg Val	528
cgc	ggc	aag	acg	gcg	gcg	ctg	ttt	tcc	gcc	gcg	acc	gag	gtg	ggc	ggc	576

G1152W0\_Sequence\_ST25. txt

Arg	Gly	Lys	Thr 180	Ala	Ala	Leu	Phe	Ser 185	Ala	Ala	Thr	Glu	Val 190	Gly	Gly		
gtc	atc	gcc	ggt	gcg	ccc	gac	gat	cag	gtg	cag	gcg	ctg	ttc	gat	tac		624
Val	Ile	Ala 195	Gly	Ala	Pro	Asp	Asp 200	Gln	Val	Gln	Ala	Leu 205	Phe	Asp	Tyr		
ggc	gac	gcc	ctg	ggc	atc	agc	ttc	cag	att	gtc	gac	gac	ctg	ctg	gat		672
Gly	Asp 210	Ala	Leu	Gly	Ile	Ser 215	Phe	Gln	Ile	Val	Asp 220	Asp	Leu	Leu	Asp		
tac	ggc	ggc	gcg	acc	gag	acg	atc	ggc	aag	aac	gtg	ggc	gac	gat	ttc		720
Tyr 225	Gly	Gly	Ala	Thr	Glu 230	Thr	Ile	Gly	Lys	Asn 235	Val	Gly	Asp	Asp	Phe 240		
cgc	gaa	cgc	aag	ctg	acc	ctg	ccg	gtc	atc	aag	gcc	atc	gcc	aag	gcc		768
Arg	Glu	Arg	Lys	Leu 245	Thr	Leu	Pro	Val	Ile 250	Lys	Ala	Ile	Ala	Lys 255	Ala		
gac	gcc	gag	gaa	cgc	gcc	ttc	tgg	acc	cgc	acc	atc	gag	gcg	ggc	gac		816
Asp	Ala	Glu 260	Glu	Arg	Ala	Phe	Trp 265	Thr	Arg	Thr	Ile	Glu	Ala 270	Gly	Asp		
cag	cgc	gac	ggc	gac	ctg	gag	cac	gcg	ctg	tgc	ctg	ctg	gcc	cgt	cac		864
Gln	Arg	Asp 275	Gly	Asp	Leu	Glu	His 280	Ala	Leu	Ser	Leu 285	Leu	Ala 285	Arg	His		
ggt	gcg	atg	gag	gcc	gcg	cgc	gcc	gat	gcg	ctg	gcc	cat	gcc	gca	cgg		912
Gly	Ala 290	Met	Glu	Ala	Ala	Arg 295	Ala	Asp	Ala	Leu	Ala 300	His	Ala	Ala	Arg		
gtc	cgc	gcg	gcg	ctg	cag	gtg	ctg	ccc	gcg	cat	ccg	att	cgc	gac	atg		960
Val 305	Arg	Ala	Ala	Leu	Gln 310	Val	Leu	Pro	Ala	His 315	Pro	Ile	Arg	Asp	Met 320		
ctg	gcc	gac	ctc	gcg	gat	ttc	gtc	gtc	agc	cgc	gtg	gcc	tga				1002
Leu	Ala	Asp	Leu	Ala 325	Asp	Phe	Val	Val	Ser 330	Arg	Val	Ala					

<210> 8  
 <211> 333  
 <212> PRT  
 <213> Paracoccus carotini faciens  
 <400> 8

Met	Thr	Val	Gln	Asp	Asn	Val	Arg	Lys	Pro	Met	Asp	Arg	Leu	Ser	Glu	
1				5					10					15		
Ala	Leu	Thr	Ala	Glu	Met	Glu	Ala	Val	Asn	Ala	Leu	Ile	Arg	Asp	Arg	
			20					25					30			
Met	Ala	Ser	Arg	His	Ala	Pro	Arg	Ile	Pro	Glu	Val	Thr	Ala	His	Leu	
		35					40					45				
Ile	Glu	Ala	Gly	Gly	Lys	Arg	Leu	Arg	Pro	Met	Leu	Thr	Leu	Ala	Ala	
	50					55					60					
Ala	Lys	Leu	Leu	Gly	Tyr	Pro	Gly	Pro	Trp	His	Val	His	Leu	Ala	Ala	
65					70				75						80	
Thr	Val	Glu	Phe	Ile	His	Thr	Ala	Thr	Leu	Leu	His	Asp	Asp	Val	Val	

Asp Glu Ser Ala Gln Arg Arg Gly Arg Pro Thr Ala Asn Leu Leu Trp  
 100 105 110  
 Asp Asn Lys Ser Ser Val Leu Val Gly Asp Tyr Leu Phe Ala Arg Ser  
 115 120 125  
 Phe Gln Leu Met Val Glu Pro Gly Asn Leu Arg Thr Leu Glu Ile Leu  
 130 135 140  
 Ala Asn Ala Ser Ala Thr Ile Ala Glu Gly Glu Val Leu Gln Leu Thr  
 145 150 155 160  
 Ala Ala Gln Asp Leu Ala Thr Asp Glu Ser Val Tyr Leu Gln Val Val  
 165 170 175  
 Arg Gly Lys Thr Ala Ala Leu Phe Ser Ala Ala Thr Glu Val Gly Gly  
 180 185 190  
 Val Ile Ala Gly Ala Pro Asp Asp Gln Val Gln Ala Leu Phe Asp Tyr  
 195 200 205  
 Gly Asp Ala Leu Gly Ile Ser Phe Gln Ile Val Asp Asp Leu Leu Asp  
 210 215 220  
 Tyr Gly Gly Ala Thr Glu Thr Ile Gly Lys Asn Val Gly Asp Asp Phe  
 225 230 235 240  
 Arg Glu Arg Lys Leu Thr Leu Pro Val Ile Lys Ala Ile Ala Lys Ala  
 245 250 255  
 Asp Ala Glu Glu Arg Ala Phe Trp Thr Arg Thr Ile Glu Ala Gly Asp  
 260 265 270  
 Gln Arg Asp Gly Asp Leu Glu His Ala Leu Ser Leu Leu Ala Arg His  
 275 280 285  
 Gly Ala Met Glu Ala Ala Arg Ala Asp Ala Leu Ala His Ala Ala Arg  
 290 295 300  
 Val Arg Ala Ala Leu Gln Val Leu Pro Ala His Pro Ile Arg Asp Met  
 305 310 315 320  
 Leu Ala Asp Leu Ala Asp Phe Val Val Ser Arg Val Ala  
 325 330

<210> 9  
 <211> 1452  
 <212> DNA  
 <213> Paracoccus carotini faciens

G1152W0\_Sequence\_ST25. txt

```

<220>
<221>  mi sc_feature
<222>  (1350)..(1350)
<223>  n i s a, c, g, or t

<400>  9
agtttgatcc tggctcagaa cgaacgctgg cggcaggctt aacacatgca agtcgagcga      60
gaccttcggg tctagcggcg gacgggtgag taacgcgtgg gaacgtgccc ttctctacgg      120
aatagccccg ggaaactggg agtaataccg tatacgccct ttgggggaaa gatttatcgg      180
agaaggatcg gcccgcgttg gattaggtag ttggtggggt aatggcccac caagccgacg      240
atccatagct ggtttgagag gatgatcagc cactctggga ctgagacacg gccagactc      300
ctacgggagg cagcagtggg gaatcttaga caatgggggc aaccctgata tagccatgcc      360
gcgtgagtga tgaaggcctt agggttgtaa agctctttca gctgggaaga taatgacggt      420
accagcagaa gaagccccgg ctaactccgt gccagcagcc gcggtataac ggagggggct      480
agcgttggtt ggaattactg ggcgtaaagc gcacgtaggc ggactggaaa gtcagagggtg      540
aatcccagg gctcaacctt ggaactgcct ttgaaactat cagtctggag ttcgagagag      600
gtgagtggaa ttccgagtgt agaggtgaaa ttcgtagata ttcggaggaa caccagtggc      660
gaaggcggct cactggctcg atactgacgc tgaggtgcga aagcgtgggg agcaaacagg      720
attagatacc ctggtagtcc acgccgtaaa cgatgaatgc cagacgtcgg caagcatgct      780
tgtcgggtgc acacctaacg gattaagcat tccgcctggg gagtacggtc gcaagattaa      840
aactcaaagg aattgacggg ggccccgaca agcgggtggag catgtggttt aattcgaagc      900
aacgcgcaga accttaccaa cccttgacat ggcaggaccg ctggagagat tcagctttct      960
cgtaagagac ctgcacacag gtgctgcatg gctgtcgtca gctcgtgtcg tgagatgttc     1020
ggttaagtcc ggcaacgagc gcaaccacg tccctagtgt ccagcaattc agttgggaac     1080
tctatggaaa ctgccgatga taagtcggag gaaggtgtgg atgacgtcaa gtcctcatgg     1140
gccttacggg ttgggctaca cacgtgctac aatggtggtg acagtggggt aatccccaaa     1200
agccatctca gttcggattg tcctctgcaa ctcgagggca tgaagttgga atcgctagta     1260
atcgcggaac agcatgccgc ggtgaatacg ttccggggcc ttgtacacac cgcccgtcac     1320
accatgggag ttggttctac ccgacgacgn tgcgctaacc ttcggggggc aggcggccac     1380
ggtaggatca gcgactgggg tgaagtcgta acaaggtagc cgtaggggaa cctgcggtcg     1440
gatcacctcc tt                                                                1452

```